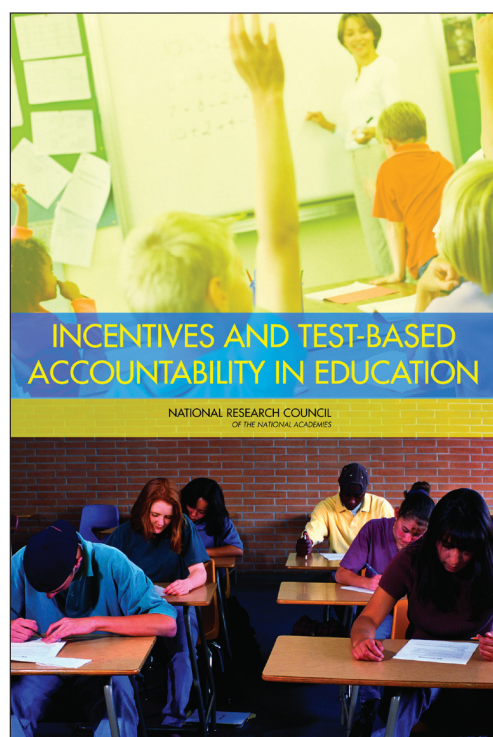


REPORT BRIEF • APRIL 2012

INCENTIVES AND TEST-BASED ACCOUNTABILITY IN EDUCATION



The past three decades have seen increasing efforts by federal and state governments to hold students, teachers, and schools accountable for how much students learn. The expectation is that such accountability will lead to improvements in education. The accountability systems often use incentives for students' performance on large-scale standardized tests.

Incentive programs offer rewards or impose sanctions on schools, teachers, or students on the basis of how students perform on standardized tests. The incentive programs in widespread use today include:

- The No Child Left Behind Act of 2001, which places sanctions on schools, such as required program changes or restructuring, if students do not show adequate yearly progress on standardized tests of reading and math.
- State high school exit exams, which require students to pass tests in multiple subjects before they can graduate (currently used in 25 states).
- Programs that tie teacher salary increases to their students' gains on standardized tests.

Although such incentive programs dominate current discussions about education policy, there has been little careful examination of what impact they have on student learning. Thus, it is important to examine the results from studies of incentives and determine whether they are contributing to the desired outcome and what policy changes may be needed.

A committee of the National Research Council examined and synthesized research on how various types of incentives—those targeted to schools, teachers, and students themselves—affect student learning. The study committee also examined the economic and psychological literature for insights that could inform modifications of existing programs or the design of new ones.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

CONSIDERING ONLY HIGH-QUALITY STUDIES

The committee adopted a rigorous approach to evaluating existing research evidence about incentive programs. The committee's report, *Incentives and Test-Based Accountability in Education*, only considered studies that allow researchers to draw causal conclusions about the effects of test-based incentive programs. This means that studies had to include a comparison group who participated in the same program but without incentives. The report did not include studies if the comparison group may have self-selected into either group.

Using these criteria, the committee examined studies of 15 incentive programs in the United States and abroad. The programs included those that imposed sanctions on schools or students as well as those that offered rewards to students or teachers. The programs studied ranged in size from national

policies affecting millions of students to careful experiments affecting only a few schools.

Another important factor that the committee examined was whether achievement results were reported on a "low-stakes" test in addition to the test with incentives attached to it (a "high-stakes" test). Attaching incentives to test scores can encourage teachers to focus narrowly on only the material that will be tested—in other words, to "teach to the test." As a result, students' knowledge of the part of the subject matter that appears on the test may increase, even as their understanding of the other parts of the subject matter (the untested portions) may stay the same or even decrease. Thus, their scores may be artificially "inflated" because the score increases reflect only part of the material that the students should know about the subject. To control for possible score inflation, researchers need to look at the effects of incentive programs on student scores on low-stakes tests, such as the National Assessment of Education-

TYPE OF INCENTIVE PROGRAM	IMPACTS ON TEST SCORES: OVERALL EFFECT SIZE ¹	CHANGE IN HIGH SCHOOL GRADUATION OR CERTIFICATION RATE ²
Imposes Sanctions on Schools: NCLB and Its Predecessors	0.08, 0.08, 0.12 ³ , 0.22 ⁴ , 0.04	
Imposes Sanctions on Students: High School Exit Exams	0.00	-2.1%, -0.6% ⁵
Offers Cash Rewards to Teachers: United States	(0.04), (-0.02), 0.01	0.9%
Offers Cash Rewards to Teachers: Foreign Countries	(0.19), (0.11), (0.19), 0.01	2.2%
Offers Cash Rewards to Students: United States	0.01, (0.06)	0.9% ⁶
Offers Cash Rewards to Students: Foreign Countries	0.19	5.4%

¹Effect size is presented in standard deviation units. Effect sizes in parentheses are derived from the high-stake test that is used in the incentive program. Studies are listed within each group in the order presented in Table 4-2 of the report.

²Studies are listed within each group in the order presented in Table 4-3 of the report.

³Study omits 8th grade reading.

⁴Study omits 8th grade reading; uses comparison to private schools during period of fluctuating enrollment.

⁵Study treats GED as equivalent to high school diploma.

⁶Same study as listed above under Offers Cash Rewards to Teachers: United States.

al Progress. Those tests are more likely to represent student knowledge of the full subject area.

In conducting its evaluation, the committee focused primarily on studies that based their assessments on low-stakes tests. For studies that provided results only for changes in the high-stakes tests, it is important to keep in mind that the results are probably affected by some amount of score inflation.

The table includes both types of studies—ones for which the impact on test scores is measured using low-stakes tests and ones for which the impact is measured using high-stakes tests. The studies that used high-stakes tests are given in parentheses to indicate that the scores are probably inflated by the incentives and thus should be given less weight because the true impacts on student learning are probably lower than the effect sizes suggest.

EFFECTS ON STUDENT LEARNING HAVE BEEN SMALL, VARIABLE

Based on its synthesis of these studies, the committee concluded that **test-based incentive programs, as designed and implemented in the programs that have been carefully studied, have not increased student achievement enough to bring the United States close to the levels of the highest achieving countries.** When evaluated using relevant low-stakes tests, the overall effects on achievement tend to be small and are effectively zero for a number of programs.

Even when evaluated using the high-stakes tests, for which the results may be inflated, a number of programs show only small increases in achievement. Of the programs studied, school-level incentive programs like those of No Child Left Behind produce some of the larger estimates of achievement effects, with effect sizes around 0.08 standard deviations. This is the equivalent of moving a child currently performing at the 50th percentile to the 53rd percentile. For comparison, raising student performance in the United States to the level of the highest performing nations would require a gain equivalent of a student climbing from the 50th to the

Insights From Psychological Research

Although much of the research on how incentives affect behavior comes from the field of economics, research from psychology offers insights too—among them the counterintuitive finding that tangible rewards can substantially undermine a person's internal motivation to perform a task. How rewards are framed matters: Several studies have shown that rewards or praise that signals competence or are purely informational ("you did that job well") tends to increase internal motivation, while that perceived as controlling or pressuring ("good, you did just as you should have") tends to decrease it. These insights from psychology are rarely reflected in the design of test-based incentive programs or in the research about their effectiveness.

84th percentile. However, although an effect size of 0.08 is small, few other education interventions have shown greater gains.

The studies also reveal that **high school exit exam programs, as currently implemented in the United States, decrease high school graduation rates without increasing achievement.** The best available estimate suggests a decrease of 2 percentage points in the graduation rate when averaged over the population.

EXPERIMENTING—AND EVALUATING—WITH CARE

The committee's evaluation suggests that despite using test-based incentives for several decades, policymakers and educators do not yet know how to design and implement them to consistently and substantially increase achievement and to improve education. Substantial further research is needed to understand how they can be used successfully.

Policymakers should support the development and evaluation of new models that use test-based incentives in more sophisticated ways as one aspect of a richer accountability system. However, given the modest and variable benefits conferred by test-based incentives so far, researchers and educators need to carefully study them and their uses to determine what works and what does not.

New test-based incentive programs should be designed and evaluated in ways that shed light on particular factors that may influence their effectiveness. Which, if any, programs are effective may depend on:

- who is targeted for incentives—teachers, students, or schools;
- what tests or performance measures are used;
- what consequences are used;

- whether additional support is provided to teachers, schools, and students in their efforts to improve; and
- how incentives are framed and communicated since this has been shown to affect the motivation of students and teachers (see sidebar).

Research should not only document the way a program is structured, but also assess a full range of outcomes. Scores on low-stakes tests are one such outcome, as are later performance in education or work and attitudes toward education.

Moreover, continued experimentation with test-based incentives should not displace investment in developing other aspects of the education system—such as improvements in curricula and instructional methods—that are important complements to any incentives.

COMMITTEE ON INCENTIVES AND TEST-BASED ACCOUNTABILITY IN PUBLIC EDUCATION

MICHAEL HOUT (*Chair*), Department of Sociology, University of California, Berkeley; **DAN ARIELY**, Fuqua School of Business, Center for Cognitive Neuroscience, and School of Medicine, Duke University; **GEORGE P. BAKER III**, Harvard Business School; **HENRY BRAUN**, Lynch School of Education, Boston College; **ANTHONY S. BRYK**, Carnegie Foundation for the Advancement of Teaching (*until 2008*); **EDWARD L. DECI**, Department of Psychology, University of Rochester; **CHRISTOPHER F. EDLEY, JR.**, School of Law, University of California, Berkeley; **GENO FLORES**, California Department of Education; **CAROLYN J. HEINRICH**, LaFollette School of Public Affairs, University of Wisconsin–Madison; **PAUL HILL**, School of Public Affairs, University of Washington; **THOMAS J. KANE**, Graduate School of Education, Harvard University, and Bill & Melinda Gates Foundation, Seattle, Washington (*until February 2009*); **DANIEL M. KORETZ**, Graduate School of Education, Harvard University; **KEVIN LANG**, Department of Economics, Boston University; **SUSANNA LOEB**, School of Education, Stanford University; **MICHAEL LOVAGLIA**, Department of Sociology, University of Iowa, Iowa City; **LORRIE A. SHEPARD**, School of Education, University of Colorado, Boulder; **BRIAN STECHER**, RAND Corporation, Santa Monica, California; **STUART ELLIOTT**, *Study Director*

FOR MORE INFORMATION... This brief was prepared by the Board on Testing and Assessment of the Division of Behavioral and Social Sciences and Education based on the report *Incentives and Test-Based Accountability in Education*. The study was sponsored by the Carnegie Corporation of New York and the William and Flora Hewlett Foundation. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not reflect those of the sponsoring organizations. Copies of the report are available from the National Academies Press, 500 Fifth Street, N.W., Washington, DC 20001; (800) 624-6242; <http://www.nap.edu>.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The nation turns to the National Academies—National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council—for independent, objective advice on issues that affect people's lives worldwide.

www.national-academies.org