

Policy Analysis

No. 661

February 17, 2010

Behind the Curtain *Assessing the Case for National Curriculum Standards*

by Neal McCluskey

Executive Summary

The argument for national curriculum standards sounds simple: set high standards, make all schools meet them, and watch American students achieve at high levels. It is straightforward and compelling, and it is driving a sea change in American education policy.

Unfortunately, setting high standards and getting American students to hit them is extremely difficult. Politically powerful interest groups must be overcome. Crippling conflicts between different religious, ethnic, and ideological factions must be avoided. And a culture that is generally averse to an intense focus on academics must be transformed. These challenges help to explain why the research on national standards is both very limited and inconclusive.

But what if the research were to clearly show that having national standards leads to superior performance on international tests? Still, there would not be compelling evidence that national standards produce optimal outcomes; economic growth, as well as personal fulfillment, could very well require an education focused on much more than just high test scores.

It appears that the route to successful education goes in the opposite direction of national standards; it goes toward universal school choice. Only a free market can produce the mix of high standards, accountability, and flexibility that is essential to achieving optimal educational outcomes.

*Neal McCluskey is associate director of the Cato Institute's Center for Educational Freedom and author of the book *Feds in the Classroom: How Big Government Corrupts, Cripples, and Compromises American Education* (Lanham, MD: Rowman and Littlefield, 2007).*

The argument seems compelling: if you want all students to do well you simply need to set a high national bar. Political and educational reality, however, are hardly so simple.

Introduction

Since at least the publication of *A Nation at Risk* in 1983, theories have abounded about what ails American elementary and secondary education.¹ One of the major contentions has been that the nation's public schools simply have not been held to high enough standards, that students are not required to clear sufficiently high bars of knowledge and understanding.

To remedy this situation, advocates of "standards-based reform" have tried to prod states to set high standards for themselves. More recently, they have championed federal efforts meant to force states to establish standards and tests. The No Child Left Behind Act requires all states to set standards in reading, mathematics, and science, test student mastery of those standards, and get all students to reading and mathematics "proficiency" by 2014.

The effect of these reforms has been spotty at best. On their own, states have set standards of widely varying—but generally low—quality. NCLB arguably made matters worse, giving states powerful incentives to set low standards and simply label poor performance "proficient." That has helped to keep states and districts out of trouble under the law, but has defeated NCLB's supposed intent: to push standards and performance to much higher levels.

Confronted with implacable state-level resistance to high standards, as well as the perverse incentives of NCLB, many standards-based reformers have moved to a new model: national standards. In defense of their proposal, some advocates argue that, in a "flat world" of global competition, it is nonsensical for a single nation to have 50 standards. Two plus two equals four whether a child lives in California, Iowa, or New York. Moreover, they argue that almost every country that outpaces the United States on international assessments has national standards. Finally, they assert that national standards will introduce desperately needed trans-

parency in American elementary and secondary education. If all states have the same standards, then the determination of a child's proficiency will not be based on where he lives, but what he actually knows.

These arguments seem compelling, centering on the idea that if you want all American students to do well you simply need to set a high national bar and get them over it. Political and educational reality, however, are hardly so simple. Examining those realities reveals why states have proven so unfriendly toward high standards on their own, and why the empirical evidence on national and sub-national standards is essentially inconclusive.

What Is Happening Now

Because of growing frustration with NCLB and escalating concerns about the nation's economic competitiveness, momentum to establish national curricular standards has markedly increased over the last few years. Several organizations have published reports calling for national standards, and many high-profile education leaders have advocated them.² Recommendations about the forms that national-standards regimes should take have varied markedly—for instance, common standards voluntarily adopted by states; standards adopted with monetary incentives from Washington; or standards with or without accompanying tests—but the unifying thread has been to have a single standard applicable nationwide.

The most concrete action has been undertaken by the National Governors Association and the Council of Chief State School Officers, which launched their Common Core State Standards Initiative in April 2009. As of September 2009, 51 states and territories had agreed to support the development of "internationally benchmarked" English and mathematics standards. If such standards are completed and states choose to adopt them, states will have to agree to make the standards at least 85 percent of their over-

all English and mathematics standards. The CCSSI released draft “college and career readiness standards” in September 2009, and draft standards for grades K–12 are scheduled to be published in early 2010.³

President Obama’s administration has offered both strong rhetorical and monetary support for adopting CCSSI standards. Well before there were publicly available drafts of the standards, U.S. Secretary of Education Arne Duncan suggested that states that adopt them would put themselves in a better position to get part of the \$4.35 billion “Race to the Top Fund”—a chunk of the massive 2009 “stimulus” bill—that he controls. He also announced that up to \$350 million of the fund would be used to develop assessments aligned with the standards.⁴ And this is to say nothing of how the standards would be integrated into reauthorization of NCLB—the core of federal K–12 policy—a process that is currently three years overdue.

Why National Standards?

National-standards advocates offer several theoretical arguments for their proposals, but almost all boil down to the same proposition: setting clear, high standards would create a simpler, more coherent, and ultimately more effective education system. As a result of having uniform national standards, the argument goes, all students, schools, districts, and states would be evaluated on the same metric, and the ability to hide failings by altering standards or tests, as is pervasive under NCLB, would disappear. A second major argument furnished by proponents is that in the modern world, with its increasingly integrated, globalized economy, it makes no sense to continue having multiple standards within one country.

Those arguments notwithstanding, it is possible that standards advocates could have been satisfied with state-level control. Until recently, states have, in fact, been the primary focus of standards-based reforms, and it is generally acknowledged that most authority over education resides at the state level.⁵

Similarly, NCLB was intended to push states to set their own standards and tests, not to have them designated nationally. But “standards and accountability” advocates have become increasingly disappointed by states’ standards—and efforts to have all children meet them—leading many to conclude that states simply will not force themselves to perform. As analysts from the pro-national standards Thomas B. Fordham Foundation have written:

The state standards movement has been in place for almost fifteen years. For almost ten of those years, we . . . have reviewed the quality of state standards. Most were mediocre-to-bad ten years ago, and most are mediocre-to-bad today. They are generally vague, politicized, and awash in wrongheaded fads and nostrums. With a few exceptions, states have been incapable (or unwilling) to set clear, coherent standards, and develop tests with a rigorous definition of proficiency. By our lights, you can count on one hand the number of states with clear proficiency standards in reading and math and expectations even approaching those of the National Assessment of Educational Progress.⁶

States, as the passage suggests, have tended to set standards low, at least compared to what groups like the Fordham Foundation hope for. It is a situation that has been exacerbated by NCLB, which leaves it to states to write their own standards, prepare and administer their own tests, and define proficiency for themselves. The result has been that most states, to comply with the letter of the law while keeping out of trouble, have set their standards very low.

A 2009 analysis that correlated state test scores with performance levels on the National Assessment of Educational Progress—a regime of federal exams used to gauge American students’ knowledge and skills in several subjects—made clear how low and variable state standards are. On fourth-grade mathematics,

Secretary of Education Duncan suggested that states that adopt standards would be in a better position to get “Race to the Top” funds.

Most appeals have been couched in superficial assertions about a modern nation needing uniform standards and have stopped there.

only one state out of the 48 with available data had set its proficiency level on par with NAEP's proficiency level; the rest had set it either at or below NAEP's "basic" level. This was repeated on eighth-grade mathematics, where only two of 47 states set their proficiency level at NAEP proficiency. On fourth-grade reading, zero states out of 48 had set their proficiency levels equivalent to NAEP's; a large majority had set it below NAEP's basic level. The results were just slightly better in eighth-grade reading; again, no states set their proficiency on par with NAEP, though more set it equivalent to the basic level.⁷

States' failure to independently set high standards, and the often huge difference in the definition of "proficiency" between states, are pushing standards-based reformers to focus on the national level. By ending what Secretary of Education Duncan characterizes as the craziness "of having 50 states designing their own standards," states would lose their ability to "game" accountability by making their standards easier to meet.⁸

In addition to improving clarity and coherence for state-to-state and district-to-district comparisons, many in the national-standards camp argue that we must be able to compare American students internationally. To do that, we must benchmark the achievement of all students in all states against international standards. As argued in a recent report from the National Governors Association, the Council of Chief State School Officers, and Achieve, Inc., "if state leaders want to ensure that their citizens and their economies remain competitive, they must look beyond America's borders and benchmark their education systems with the best in the world."⁹ To compete in an integrated world economy, they argue, we must hold all students to "world-class" standards.

The First Blush Fades: The Weak Theoretical Case for National Standards

On a very basic level, the argument appears to make sense: set national standards, hold

schools accountable for meeting them, and educational excellence will ensue. One problem with this logic, however, is that it fails to tackle the bedrock question of how to design, implement, and enforce *high* standards. After all, it is getting high standards for all children—not national standards *per se*—that is the ultimate goal of standards-based reformers. But as constant state-level failures to create and meet rigorous standards have made clear, getting high standards is a very challenging undertaking. Perhaps because of that, most appeals for national standards have been couched in superficial assertions about a modern nation needing uniform standards—or algebra being the same in Maine as it is in Mississippi—and have stopped there.

That said, a few analysts have attempted to give more rigorous theoretical arguments for centralized standards and testing—though not national standards alone—than have its leading proponents.

Economist Robert Costrell has mathematically modeled the potential effects of centralized education standards. Assuming that (a) the standard is "defined as the required level of proficiency for a binary credential, such as a high-school diploma"; (b) centralized standards-setters are motivated to "maximize their conception of social welfare"; (c) "utility-maximizing students choose whether to meet the standard"; and (d) all districts are identical, Costrell asserts that standards will be more demanding if they are set at a higher, more central level than the district. He argues that in a decentralized system, without a way to compare districts or schools, districts would gain no advantage in the marketplace by setting higher standards because high school diplomas would not be valued differently by employers.¹⁰ In a subsequent analysis, Costrell adds that a hybrid system of centralized minimum standards and a decentralized ability to set higher standards would maximize social welfare, leading to an optimal societal combination of leisure and income.¹¹

Of course, modeling is not the same as testing with real-world data, so Costrell's analysis is primarily theoretical. More importantly,

some of his foundational assumptions are faulty. Costrell is incorrect, for instance, in assuming that employers, as well as colleges, cannot and do not compare districts or job applicants' skills in a decentralized system. In a recent survey of over 400 employers in the United States, almost 45 percent of respondents reported that they tested or screened "recent high school and college graduate applicants to determine proficiency in some specific basic knowledge/skills (i.e., *Math, Reading, Writing, Spoken English, or Other*)."¹² Similarly, while the college admissions process is often opaque and numerous variables are taken into account for each applicant, some colleges do systematically track the performance of specific high schools' graduates to assess the standards and performance of those schools (and, by extension, their districts).¹³

What appears to be missing is not an ability of employers and colleges to differentiate between districts, but an incentive for districts to change. Whether they perform well or not, districts typically receive uninterrupted taxpayer funding. The only methods parents and taxpayers have for punishing unsatisfactory districts, as a result, is moving to better districts or changing authority through school-board elections. But neither of these mechanisms can exert much pressure. For the former to work, dissatisfied constituents must move to a better district—a costly and inefficient way to exercise choice. And the latter?

Political realities make it very difficult to use school-board elections to induce change. Consider, for instance, the politicians who would adopt standards. These actors are as self-interested as anyone else, and as such their first concern is not to maximize student achievement or social welfare, but to win reelection. And which groups involved in education are most likely to help the politicians achieve that goal? While parents and taxpayers are certainly numerous, few are single-issue education voters. Most are either concerned with a variety of political issues or are not very politically engaged. In contrast, teachers and administrators whose livelihoods depend on public schooling are highly motivated to focus

on education, and so exert outsized power over politicians on education issues. And because employees are naturally averse to management raising the standards for their performance, it is unsurprising that teachers' unions and administrators' associations use their political influence—which goes far beyond just voting—to keep standards low.

In looking past the inherent standards-minimizing pressure of democratic control, Costrell is joined by historian Diane Ravitch. Ravitch asserts in *National Standards in American Education: A Citizen's Guide* that clear standards are necessary so that everyone involved in the education process will know what their end goals and measures of success are. She also argues that national standards must be established by some government entity lest textbook and test-writing companies set standards by default. Finally, she contends that national standards would help eliminate discriminatory expectations for students based on race or income level.¹⁴ She writes all this, despite acknowledging the huge strength of political forces arrayed against high standards. As she wrote of President Clinton's failed national standards effort:

I came to realize even as I was writing that the seemingly straightforward idea of national standards was falling victim to a tortuous political process. Other agendas became attached to the basic idea of standards, as legislators and lobbyists from various interest groups saw an opportunity to hang their favorite causes onto the legislation. The measure placed unnecessary restrictions on how states could use their tests; dictated to states the composition of their standards-setting body, ensuring the representation of every professional group and mandating affirmative-action criteria for selection; insisted on vague and controversial opportunity-to-learn standards; and required that any state educational plan encompass all social services, not just academic standards. . . . Yet I continue to believe that the idea of

What is missing is not an ability to differentiate between districts, but an incentive for districts to change.

The asymmetry of political influence between the public and public-school employees is manifested in the physical presence of education groups in Washington.

national standards has remarkable validity.¹⁵

The asymmetry of political influence between the public and public-school employees is manifested in the physical presence of education groups in Washington, DC, and state capitals. Just in and around the District of Columbia, the National Education Association has a headquarters building, and in fiscal year 2008 employed 896 people on its national staff; the American Federation of Teachers has a headquarters, and in FY 2009 employed 409 staffers; and groups such as the American Association of School Administrators, National Association of Elementary School Principals, and the Council of Chief State School Officers also have headquarters.¹⁶ In contrast, the only group ostensibly representing parents, the National Parent Teacher Association, is headquartered in Chicago and has only its lobbying arm in Washington. And the PTA is hardly representative of most parents; it is largely controlled by the NEA and has suffered years of declining membership.¹⁷

These groups, their state affiliates, and their members have been very influential in fighting efforts to implement rigorous standards and accountability. In FY 2008, the NEA reported spending almost \$30 million on “political activities and lobbying,” as well as delivering over \$82 million in “contributions, gifts, and grants” to such groups as the Democratic Leadership Council and the Ballot Initiative Strategy Center.¹⁸ The latter describes itself as the “nerve center for progressive ballot initiative campaigns across the country.”¹⁹ In FY 2009, the AFT spent over \$24 million on politics and lobbying and made about \$6 million in gifts. Included among the AFT’s gift recipients was the Economic Policy Institute, whose “Broader, Bolder” accountability initiative proposes that states undertake “qualitative evaluation of school quality and do not rely primarily on standardized test scores to judge the success of schools.”²⁰

Education policy is still, though, primarily made below the federal level, and state and local NEA and AFT affiliates are both numer-

ous and well-funded. In FY 2008, Illinois’ NEA affiliate spent more than \$1 million on lobbying and other overtly political activities. In roughly that same time, the New York State United Teachers—which is affiliated with both the NEA and AFT—spent about \$4.7 million. The Ohio Education Association expended almost \$3.4 million. The Michigan Education Association spent about \$2.5 million. In the previous year, the Pennsylvania State Education Association expended nearly \$1.9 million on political activities and \$372,713 on gifts, including \$5,000 to Fairtest, an organization that opposes the use of “test scores to make critical educational decisions about students or schools,” and on its website features endorsements from unions around the country.²¹

Going beyond teachers’ unions, groups representing numerous other public school employees actively oppose connecting standards to strong accountability based on test performance. In 2000, Paul Houston, executive director of the American Association of School Administrators, stated that “testing should be a *part* of how schools measure student performance; however, educating students . . . cannot be measured by *one* test alone.”²² In November 2006, the executive directors of the National Association of Elementary School Principals and the National Association of Secondary School Principals issued an open letter to Congress arguing that school and student progress must be measured “based on the results of multiple assessments and multiple opportunities to retake the test.”²³ Finally, in 2000, the National Council of Teachers of English published a resolution stating that “the use of any single test in making important decisions . . . is educationally unsound and unethical.”²⁴

In addition to powerful resistance from special interests, when accountability measures threaten to keep poorly performing children back a grade or from graduating high school, significant pressure to keep standards low or easily evaded comes from students, families, and sympathetic members of the public. In 2007, for instance, when it appeared

that a large percentage of students would not graduate, Washington State eliminated a pending high school graduation requirement that students pass the mathematics section of the Washington Assessment of Student Learning. That same year, the Maryland State School Board approved an alternate evaluation—a project instead of a test—to enable those who had failed the High School Assessment to graduate.

In the end, all of these political forces have rendered state standards an unappetizing hash. In 2006, analysts at the Thomas B. Fordham Foundation gave standards across the 50 states an average grade of C-minus. Only three states—California, Indiana, and Massachusetts—received As. The problem, they acknowledged, has been that the political system is stacked against high standards and tough accountability. Nonetheless, seemingly because they could think of no other option, the Fordham writers endorsed national standards:

We understand that national standards would face the same perils as state standards. If written by committee, or turned over to K-12 interest groups, they could turn out to be vague, politically correct, encyclopedic, and/or fuzzy. If linked with real consequences for schools, they could be pressured downward. They could even wind up doing more harm than good.

But if done right, they could finally put the entire country on the sturdy path of standards-based reform. And if great standards can be written in Sacramento or Indianapolis or Boston, perhaps they could be created in Washington, DC.²⁵

Maybe It's Not the Structure, But the Culture

A major factor behind low standards and poor test results is the great power that those who would be held to high standards have over the American education system. But a general American discomfort with centralization and easily tested knowledge also plays a

part—not something, as will be discussed later, that is necessarily bad.

Many Americans simply do not subscribe to centralized government control of institutions, including education, or of test-driven, fact-and-skill-centered learning. Testament to these proclivities are the popularity of child-centered Montessori and other pedagogically progressive schools, the nation's long tradition of “local control” in education, and, arguably, the general unpopularity of the No Child Left Behind Act.²⁶ It is, in fact, quite possible that centralized control of education, desire to perform well in easily tested subjects and on standardized tests, and focusing on academic pursuits at all are broadly at odds with American culture. Centralized control of education, for instance, goes against the grain of what sociologist Seymour Martin Lipset famously labeled “American exceptionalism”—a culture grounded in individual liberty and laissez-faire governance.²⁷ Similarly, the hard-work but anti-academic ethos of the colonist, frontiersman, and entrepreneur are central to the American narrative.

Just as culture in the United States might be generally averse to centralization and standardized education, other cultures might gravitate heavily in the opposite direction. This seems very much the case for East Asian nations, which typically dominate international tests.

Historian J. M. Roberts notes that in China “an enlarged bureaucracy was to survive many periods of disunity . . . and remained to the end one of the most striking and characteristic institutions of imperial China.” And how did one enter the bureaucracy? “The officials were in principal distinguished from the rest of society only by education.” Roberts also reports that despite “rich variety in culture and custom . . . all the East Asian peoples are similar in seeming to show great industry and enterprise, and a willingness to accept a marked subordination of the individual to the group.”²⁸

There is evidence within education that there is, indeed, a significant cultural difference between American and other students.

Many Americans simply do not subscribe to centralized government control of institutions.

On the 2007 eighth-grade Trends in International Mathematics and Science Study (TIMSS) mathematics assessment, 11 of the 12 lowest performers had national standards.

American schools and students tend to focus more on “critical thinking” and other less-concrete and measurable outcomes than mathematical and scientific skills and knowledge. Americans also tend to put much less emphasis on schooling and academics than the people of other industrialized nations, and much greater emphasis on extracurricular activities and part-time employment.²⁹

Specifically looking at elementary and secondary education, Stevenson, Chen, and Lee examined the attitudes toward education of children and parents in Minneapolis, Minnesota, Sendai, Japan, and Taipei, Taiwan, in 1980, 1984 and 1990. In all three years, despite significant publicity about the high achievement of Japanese and Chinese students and the low standing of Americans, the researchers found that Chinese and Japanese parents were much less satisfied with their schools than were their American counterparts. Similarly, they found that Chinese and Japanese parents and students stressed the importance of hard work to achieve academic success, whereas Americans attributed academic excellence much more to innate ability.³⁰

Finally, there is the difference in academic performance between Asian students in the United States and other U.S. racial/ethnic groups. Understanding that “Asian” is a designation that includes people as diverse as ethnic Japanese and Hmong, Asian students outperform all other U.S. students by sizable margins, even after adjusting for socioeconomic status, on mathematics assessments.³¹ And since all these groups are in the same education system, this strongly suggests that culture is a powerful force regardless of the educational structure.

Paltry Direct Empirical Evidence

As the Fordham Foundation admits, there is little theoretical reason to believe that national standards will be any better insulated against downward political pressures than state standards. Still, national-standards advo-

cates frequently suggest that, theory or no theory, national standards are achievable and will work. After all, as American Federation of Teachers president Randi Weingarten has written, “the countries that consistently outperform the United States on international assessments all have national standards.”³²

Unfortunately, this “all who beat us have national standards” factoid is close to the only empirical support that national-standards advocates typically offer for their cause. As a result, one cannot help but conclude, as shall soon be seen, that the scientific evidence on national standards is thin and furnishes insufficient support for the notion that strong national standards are either achievable—or desirable—in the United States.

Before addressing the empirical literature, it is worth dismissing the factoid. It is true that most nations that have outperformed the United States on such tests as the Trends in International Mathematics and Science Study and the Program for International Student Assessment have national standards, *but so do most nations that have done worse*. To illustrate, on the 2007 eighth-grade TIMSS mathematics assessment, the eight countries that outperformed the United States had national standards. But, then, so did 33 of the 39 nations that scored lower. Moreover, 11 of the 12 lowest performers had national standards.³³ When looking only at countries belonging to the Organization for Economic Cooperation and Development—generally, economically advanced nations—the same noncorrelation holds: four OECD members outperformed the United States, six did worse, and all but the United States and Australia had national standards. All of this holds true for the 2007 TIMSS eighth-grade science assessment, on which all 10 nations that outperformed the United States had national curricula—but so did 33 of the 37 lesser performers and the 9 *lowest* performers. Among OECD members, five posted better scores than the United States, five did worse, and only the United States and Australia did not have national standards.

The most recent TIMSS to include the

final year of secondary school—essentially, high school seniors—was the 1995 administration. Very few countries of the 21 participating had adequate student samples to reach firm conclusions, but the results again suggest that having national standards does not correlate strongly with performance. The United States finished poorly—fourth from last on the combined mathematics and science literacy scale—but the three nations it outperformed all had national standards. Meanwhile, among the top five finishers, three did not have national standards, including the top-finishing Netherlands.³⁴

The 2006 Programme for International Student Assessment science exam results, which tested 15-year-olds, exhibit similar patterns. In 2006, 57 nations participated in the exam, and PISA reported whether 56 had national, regional, or fully decentralized science standards and testing (France had insufficient data). Of the 27 nations with sufficient data that outperformed the United States, 17 had national standards and examinations, 3 had external standards and examinations in some regions, and 7 had no centralized standard—very much a mixed bag. The same was true for the 28 nations that did worse, though there were proportionately more countries without national standards: 12 had national standards and tests, 3 had regional, and 13 had no centralized standards and tests.³⁵

Importantly, few of the nations in the bottom half of the PISA standings were developed nations. How did the OECD-only breakdown look? Of the 19 OECD nations with available data that outpaced the United States, 11 had national standards and tests, 3 had regional, and 5 had no centralized standards. Of the 9 that did worse than the United States, 4 had national standards and tests, 1 had regional, and 4 had none. Clearly, having national standards is no guarantee of superior performance.

What about the influence of socio-economic status, which is a very powerful predictor of academic success? Of the five nations highlighted by PISA for having above-average performance in science coupled with below-

average impact of socioeconomic background, three had national standards and tests, while two—Australia and Canada—had regional standards like the United States.³⁶

All of this said, we have testing and standards data for a relatively small number of nations, and this has made it difficult to draw universally applicable conclusions from comparing nations with and without such standards. However, some recent empirical work has been interpreted as supporting national standards.

The first major such analysis was conducted by economist John H. Bishop in 1996, though, importantly, he looked at national standards for secondary students coupled with “curriculum-based external examinations” (CBEEs) that students must pass in order to graduate.³⁷ So far, such testing has not been proposed under leading national standards efforts in the United States. In addition, Bishop explored the effect of these exams for *secondary* students on the standardized testing results of 13-year-olds, probing for a systemic effect of CBEEs.

After controlling for such factors contributing to academic achievement as the number of books in students’ homes, number of siblings, and socioeconomic status in 15 countries, Bishop found that non-CBEE nations performed worse on the 1991 International Assessment of Educational Progress than CBEE nations. Non-CBEE nations had mathematics scores about two U.S. grade levels below CBEE countries, and the results were statistically significant. Science results were not statistically significant. (Bishop also included results for geography, which are not discussed here because that subject is not a focus of U.S. national standards efforts.)³⁸

There were numerous problems with Bishop’s study—indeed, Bishop himself stated that “the power” of his test was “very low.”³⁹ For one thing, only for the United States and Portugal, and only in mathematics, did Bishop find lower performances that were statistically significant for specific non-CBEE countries, and that significance was at the borderline 10-percent level.⁴⁰ Perhaps

Among the top five finishers, three did not have national standards, including the top-finishing Netherlands.

Nations with curriculum-based external examinations (CBEEs) might have predispositions that make them both more likely to have CBEEs and to do well on achievement tests.

more importantly, Bishop acknowledged that his findings could be spurious, indicating not that CBEEs drive higher achievement, but that nations with CBEEs might have cultural or ideological predispositions that make them both more likely to have CBEEs and to do well on achievement tests:

Causation is not proved . . . because other explanations for the U.S., Spanish, and Portuguese lag can no doubt be proposed. Other sources of variation in curriculum-based exams need to be analyzed. *Best of all would be studies which hold national culture constant.* [Italics added.]⁴¹

Lending credence to the possibility that culture drives both centralization and performance, Bishop found that the positive effect of being from an Asian nation was typically greater than the effect of having a CBEE.

In 1997 Bishop released a new study that included his IAEP findings. He added an analysis using 1994–1995 TIMSS mathematics and science data for 13-year-olds in 37 nations.⁴² Regressing participating nations' TIMSS scores against per capita gross domestic product, a dummy variable for East Asian nations, and a dummy variable for CBEE, Bishop found that the presence of a CBEE was associated with superior performance of one U.S. grade level in mathematics and 1.2 grade levels in science. Both findings were statistically significant, though the former was only at the 10 percent level.

Though more extensive than his 1996 work, Bishop's 1997 findings suffer from the same problems. Again, Bishop looked not at the effect of national standards, but standards coupled with high-stakes tests. And again, aside from creating a dummy variable for East Asian nations, Bishop did not control for cultural or other unobserved variations. That said, there was very suggestive evidence that culture might be a highly significant force behind both the creation of national standards and achievement: in both the IAEP and TIMSS mathematics results, being from

an East Asian nation had a greater positive impact than having a CBEE. Indeed, on TIMSS the effect was almost three times more powerful.

In a 1999 study Bishop again attempted to determine the effect of national standards and CBEEs on academic outcomes, this time looking at the effect of tests controlled at national and "provincial" levels.⁴³ The study included TIMSS data for two more nations than in 1997. Bishop found that the presence of high school CBEEs improved scores about one grade level for 13-year-olds in math and 1.3 grade levels in science. Bishop also used scores on the 1990 International Association of the Evaluation of Educational Achievement literacy test, which included 25 nations, and the 1991 IAEP tests of 15 nations. Those, too, showed a sizable CBEE advantage, though it was only statistically significant in math for the IAEP.

In the 1999 study, Bishop attempted to more directly address the question of whether culture might drive both the establishment of CBEEs and higher achievement, and potentially create the spurious impression that CBEEs improve achievement. Unfortunately, he only tried to control for culture in analyzing Canadian provinces, which is not directly applicable to *national* standards. Moreover, he argued only that because provinces with CBEEs did not have lower rates of disciplinary problems or absenteeism, they likely did not have greater inherent interest in education than non-CBEE provinces. But provinces in which citizens have a greater focus on education might demand more forcefully that schools fully report problems, and those provinces might do better academically despite having just as many kids causing trouble. That is likely one reason Bishop conceded again that studies adjusting for national culture were needed.⁴⁴

Following Bishop, economist Ludger Woessmann tackled the effect of standards coupled with high-stakes assessments. Like Bishop, Woessmann found a strong relationship between the presence of CBEEs and superior results on TIMSS for 13-year-olds.

Using 39 countries that participated in the 1995 TIMSS; 38 countries that participated in the 1999 repeat test; identifying whether the nations had CBEEs in mathematics and/or science; and controlling for 48 “family, resource, teacher, and institutional control variables,” Woessmann found that students in CBEE nations scored about 43 percent of a standard deviation better in math and almost 36 percent of a standard deviation better in science.⁴⁵

Woessmann’s findings were similar to Bishop’s, but so were his study’s shortcomings. Most importantly, Woessmann did not control culture adequately to rule it out as a major unobserved driver of both CBEEs and high achievement. Arguing that “concerns about cultural differences generally arise in cross-regional comparisons . . . but should not be as large within regions,” Woessmann created dummy variables for regions of the world to isolate effects within regions.⁴⁶ However, it is a questionable assumption that the differences among nations within a region are not nearly as great as those between regions. Looking at the variation among nations within a region suggests potentially large gulfs. Iran and Israel hardly have the same culture, but they are both in the Middle East. Japan and the Philippines are both in Asia. Spain and Denmark are both in Europe. Finally, for all intents and purposes when looking at TIMSS data, many of the “regions” identified by Woessmann are just a few nations. While his Middle East consists of 4 nations, Asia 9, Eastern Europe 13, and Western Europe 20, North America, South America, Oceania, and North Africa consist of only 2 nations, and South Africa only 1.⁴⁷

That said, in a working paper very similar to his published analysis (with similar data, controls, and findings), Woessmann listed the coefficients for each region (he did not publish them in his final version). Those coefficients indicated that there very well could be a cultural effect. Like Bishop, Woessmann’s data showed a very strong, positive Asian effect in math, with the Asia dummy variables providing a boost more than

twice as large as having a CBEE. The Eastern European dummy variables also had a sizable positive effect, coming close to being twice as large as the CBEE effect. There were no statistically significant positive regional effects found in science, but there were significant *disadvantages* in science for the Middle East and Southern and Northern Africa.⁴⁸

In 2004, German researchers Hendrik Jürges and Kerstin Schneider attempted to determine the causes of big differences in academic achievement between otherwise similar nations, and included an analysis of CBEEs. Where they differed from Woessmann and Bishop was in using only nations belonging to the OECD. In so doing, they got results very much at odds with Bishop’s and Woessmann’s: CBEEs had negligible effects on TIMSS math achievement.

Why the difference? For one thing, the authors note, Bishop included greater numbers of less-developed nations in his analysis than they did. In addition, Bishop included some highly distorting outliers, especially the CBEE-less Philippines, which Jürges and Schneider found to be “a highly influential observation.” After adjusting for outliers, Jürges and Schneider concluded that “central exit examinations do not have a significant impact on the achievement of middle-school students.”⁴⁹ Jürges and Schneider did, however, find one thing that was very similar to Bishop’s results: a highly positive effect for Asian nations, suggesting, once again, that culture could have a powerful influence on achievement.

The most recent study touting national standards comes from the Thomas B. Fordham Foundation.⁵⁰ In it, Michigan State University professor William H. Schmidt and his coauthors argue that several countries have instituted national standards—including Germany, which the paper dubiously states has a federalist tradition like the United States—and that the United States could, too. The paper offers no systematic evidence that as a result of having national standards countries with such standards outperform the United States, but asserts it

After adjusting for outliers “central exit examinations [did] not have a significant impact” on achievement.

The scientific literature on national standards simply does not demonstrate that national standards drive superior performance.

nonetheless. This despite Schmidt having conceded, when asked at a conference preceding publication of the paper what empirical evidence demonstrates the beneficial achievement effects of national standards, that “it is very difficult to establish that, because virtually everybody in the world has national standards.”⁵¹

Thin Indirect Empirical Evidence

The scientific literature on national standards is too thin to support any U.S. policy move in that direction. Moreover, what research does exist is handicapped by the potentially confounding effects of unobserved variables such as culture, and focuses not just on national standards, but on standards coupled with high stakes for students. It simply does not demonstrate that national standards drive superior performance on standardized tests. Perhaps, though, the research on sub-national standards is more illuminating.

Bishop supplements his analysis of national-level data with comparisons of academic outcomes in Canadian provinces, which, like U.S. states, have autonomy to set (or not set) their own standards. This is important because, with no national standards, Canada has done very well on international comparisons, especially PISA. In 2006, Canadian 15-year-olds finished third out of 57 nations in scientific literacy, fourth in reading, and seventh in mathematics.⁵² Of course, this does not prove that *lacking* national standards drives superior performance any more than high-performing countries having national standards proves the opposite. It is also important to remember that performance data come from specific tests, and strictly speaking demonstrate only the correlation between having standards and performance on those tests. And, as shall be discussed further, test results do not capture all, or perhaps even most, desired educational outcomes.⁵³

In his 1996 work, Bishop compared the

IAEP scores in nine Canadian provinces and the entire United States, adjusting for several variables, including the presence of a CBEE. Four provinces—Alberta, British Columbia, Newfoundland, and Quebec—had a CBEE, while the remaining five and the United States did not. Bishop again found that having a CBEE led to superior performance, on the magnitude of between two-thirds and four-fifths of a U.S. grade-level in mathematics (depending on whether some variables were counted as influenced or not influenced by CBEEs) and two-thirds of a grade-level in science.⁵⁴ Bishop reiterated the findings in his 1997 and 1999 work.

Despite his findings, Bishop’s analysis of Canadian provinces and the United States failed again to demonstrate that centralized standards drive superior academic performance. Once again, he assessed the effect not just of centralized standards, but standards attached to exams with high stakes for students. And once again, Bishop did not control for differences in culture, attitudes toward education, or other potentially confounding variables.

Along the lines of Bishop’s provincial assessment, several researchers have tried to determine the impact of state-level standards and tests on academic outcomes within the United States. The results, however, have been as problematic as the evaluations of Canadian-province and national-level standards.

Perhaps the most publicized research on state standards was a series of studies by Arizona State University professors Audrey Amrein and David Berliner. Looking at state standards and testing coupled with high-stakes for either schools, students, or both, Amrein and Berliner determined that high-stakes accountability in elementary and middle schools was associated with no systematic increase or decrease in academic achievement as measured by NAEP. At the high school level, implementation was associated with *decreasing* ACT, SAT, and Advanced Placement scores.⁵⁵

Amrein and Berliner’s results garnered significant attention, but suffered from serious

methodological flaws. University of Illinois education professor Barak Rosenshine noted, for instance, that Amrein and Berliner compared results in states with high-stakes systems against national averages—a mistake given that national averages included high-stakes states. Instead, comparing the performance of high-stakes states against low- or no-stakes states, Rosenshine found that high-stakes accountability had a moderate-to-large positive effect. He also noted that Amrein and Berliner considered smaller than national-average *increases* on state NAEP scores to be *decreases*. Looking at absolute changes, Rosenshine found that only two high-stakes states saw scores declines, versus 10 low-stakes states.⁵⁶

In response to Rosenshine, Amrein and Berliner compared high-stakes to non-high-stakes states and found that high-stakes states were outperforming low-stakes states. However, the results were only statistically significant on fourth-grade mathematics, not fourth-grade reading or eighth-grade mathematics. They also argued that students were being exempted from tests at greater rates in states with high stakes than those without, and that that was a greater explanation for rising scores than the presence of high-stakes tests.⁵⁷

Around the time of Rosenshine's exchange with Amrein and Berliner, Hoover Institution researchers Margaret Raymond and Eric Hanushek published a blistering treatment of Amrein and Berliner's study. They lodged the same complaints as Rosenshine, but also took issue with Amrein and Berliner's disregarding score increases in states in which exclusion rates increased, something Amrein and Berliner did on the premise that exclusion contaminated the results. Reassessing the findings, Raymond and Hanushek found that high-stakes accountability had a significant, positive effect.⁵⁸ However, they provided no state-by-state data to support their findings, nor did they control for unobserved variables such as culture, ideology, or potential "shocks" such as sobering reports about academic outcomes that might have changed peo-

ple's attitudes toward education.

In 2002, economists Martin Carnoy and Susanna Loeb released a study on the effect of external accountability on student outcomes. Comparing changes in the percentage of students scoring "basic" or above on NAEP to the strength of accountability (roughly, the degree to which stakes were attached to testing outcomes), the authors found that states that had implemented stronger accountability in the early 1990s saw greater fourth- and eighth-grade mathematics gains between 1996 and 2000. They also provided a glimpse into the possibility that variables such as culture or ideology might drive both centralization of accountability and achievement. Concerning centralization of standards and accountability, they reported that more governmentally centralized states, especially in the South, tended to have higher-stakes accountability, as did "more populous states with correspondingly larger absolute numbers of disadvantaged minorities, larger school systems, and larger cities."⁵⁹ Concerning achievement, they found "no relationship between test-score gains prior to implementation and the strength of accountability," hinting that a cultural or ideological proclivity toward high achievement did *not* drive scores. The period "prior to implementation," however, was just a few years—too short to reliably identify a state's culture or potential effects of changing attitudes toward education.⁶⁰

Perhaps the most important data in Carnoy and Loeb's study were not central tendencies, but graphs of state-by-state results for eighth-grade mathematics. Though "tougher" accountability (measured on a 0 to 5 scale) was correlated with better outcomes, the variation from state to state was considerable, illustrating the pitfalls of assuming that central tendencies are universally applicable. For white students, while the 14 lowest-stakes states (out of 37 total) tended to have the smallest gains, three of those states outpaced five higher-stakes states and tied five more. And the overall outcomes for states in the three highest accountability levels were heavily influenced by just two very high perform-

Outcomes for states in the highest accountability levels were heavily influenced by just two very high performers.

Except for Bavaria, all of the German states with central exit exams were in the former East Germany.

ers: Louisiana and North Carolina, which saw 15 and 14 percentage-point increases, respectively. (There were also two states in the highest accountability levels—California and New Mexico—that saw *zero* gains.) Were Louisiana and North Carolina eliminated, the change in the percentage of white students scoring “basic” for states in the top three levels would drop from 6.5 points to 5.2. Meanwhile, the average increase for states between 1 and 3 on the accountability scale was 5.4 percent. Results for black and Hispanic students were similarly volatile.⁶¹

In January 2004, again expanding on the work of Amrein and Berliner, education analyst Henry Braun examined changes in NAEP scores for states with and without high-stakes standards and tests. He found, as others had, that high-stakes states showed greater grade-level improvements on NAEP than did states with lesser stakes. However, when Braun looked at improvements for specific cohorts of students, such as between 1996 fourth-grade scores and 2000 eighth-grade scores, low-stakes states had an advantage. He also reiterated a very important caveat: “conclusions from analyses concerning the effects of testing or, more generally, accountability policies must be tentative, based as they are on highly aggregated, observational data.”⁶²

In a 2005 article, researchers Hendrik Jürges, Kerstin Schneider, and Felix Büchel attempted to determine the effect of central exit exams while better controlling for variables such as culture. They did so through a difference-in-differences framework that looked at the differential for vocational-track students between TIMSS mathematics and science scores in German states with and without centralized exit exams. This helped to estimate the effect of unobserved variables because no states with central exams required vocational-track students to take them in science, while all required mathematics. If the exams drove performance, one would expect to see bigger math-science differentials in exit exam than non-exam states.

Jürges, Schneider, and Büchel did have some concerns. First, if improving mathe-

matics skills also improves science outcomes, the effect of exit exams could be understated, with science scores keeping up with math in part because of the math exams. This did not greatly concern the authors, however, because only a small fraction of TIMSS science questions called on math skills. On the flip side, testing math could “crowd out” science instruction, overstating the effect of exit exams. In an attempt to control for this, the authors included as an independent variable the number of hours students reportedly spent studying science outside of school. Finally, it is important to keep in mind that “college-prep” students were not included in this study, limiting the applicability of its findings to those students.

The authors found that their analysis dramatically cut the “raw” exit examination advantage: from about 1.25 school years to an estimated one-third of a year, bolstering the culture theory while suggesting that exams have some independent positive effect.⁶³ The study also provided another indication that differing cultures and ideologies might affect where centralized exams would be found and, perhaps, be more effective: except for Bavaria, all of the German states with central exit exams were in the former East Germany.

In the same year that Jürges, Schneider, and Büchel’s work was published, Hanushek and Raymond attempted to identify the effect of imposing state standards and tests by comparing states’ NAEP score growth from fourth to eighth grade, which would include a single cohort of students. Looking at scores between 1992 and 2002, they found that state accountability systems had a positive effect on achievement. They also provided potential insight into the culture question with their own difference-in-differences design, which controlled for unobserved state variables “that exert a constant influence on state performance over the relevant observation period.”⁶⁴ This could have controlled for culture, but the period covered—essentially, the early 1990s is the entire no-standards period—is too short to reach any firm conclusions about

changing culture or attitudes. In addition, although the authors provided an analysis without four outliers—California, North Carolina, Texas, and Washington, DC—and found that the outliers had little effect on overall results, they offered no data broken down by state showing how individual states differed from overall averages.

The Research Summarized

Having reviewed the extant comparative research on standards set both at national and subnational levels, it is clear that the empirical support for national standards is weak. First, there simply has not been much comparative research done. Second, what little has been done has typically focused not simply on national standards, but standards coupled with high stakes for students, something not contemplated, at least publicly, under the Common Core State Standards Initiative. Finally, the research does not control adequately for unobserved variables, especially culture or changes in attitudes, that could drive both centralization and a focus on easily tested academic achievement.

Problems beyond the Weak Research Base

The first question that needs to be answered before fundamentally altering the status quo is whether a given reform will work. The existing national-standards research provides no conclusive answer. Beyond the huge deficiencies in the empirical case, however, there are other important objections to imposing national standards on American education.

Federal and Involuntary

Many national standards advocates are quick to point out that the CCSSI—the leading national standards effort—involves states cooperatively determining standards and voluntarily adopting them, not federal control.

As the CCSSI website explains: “Governors and state commissioners of education from across the country committed to joining a state-led process to develop a common core of state standards in English-language arts and mathematics for grades K–12.”⁶⁵

There is a sound political reason for emphasizing state control: citizens and state officials are loath to have standards imposed on them by Washington, and previous efforts to create standards at the federal level failed miserably.⁶⁶ Moreover, many people fear that having the federal government set standards and write tests would make it easy for special-interest groups like teacher unions and administrator associations to keep standards low. They would essentially have “one-stop shopping,” a single power center on which to focus all of their energies. As Sandy Kress, a chief architect of the No Child Left Behind Act, has stated, “the process [of setting standards] will be hijacked by [interest] groups if the process is federal.”⁶⁷

That said, despite national-standards supporters emphasizing that the CCSSI is state-led and that adoption of its standards is technically voluntary, adoption will almost certainly be de facto involuntary, and the standards themselves ultimately federal. Already, as previously noted, Secretary of Education Duncan has made clear that it would behoove states to sign on to the CCSSI if they wish to compete for a share of the \$4.35 billion “Race-to-the-Top” fund. Duncan has also said that the federal government would furnish \$350 million to develop tests tied to the standards. Finally, it would be very difficult to conceive of a reauthorized No Child Left Behind Act, especially in light of what Duncan is currently pushing, that would not mandate adoption of national standards and ultimately lead to federal control of those standards. Of course, states could “voluntarily” turn down the billions of federal dollars attached to NCLB, but even in that unlikely case there would still be absolute compulsion involved: state citizens would continue to have no choice about paying federal taxes.

Beyond the huge deficiencies in the empirical case, there are other important objections to imposing national standards.

By demanding that all states adopt specific standards, Washington would be exceeding even the unconstitutional power it has accumulated under ESEA.

Constitutionality

The federal government has only been heavily involved in education since 1965, when the Elementary and Secondary Education Act was passed, because for most of the nation's history it was understood that federal involvement would be unconstitutional. The Constitution makes no mention of "education" or "schooling" among the specific, enumerated powers it gives to the federal government, and outside of controlling the District of Columbia and military educational activities, Washington has no authority to be involved in education.⁶⁸

To justify its growing involvement, federal policymakers have typically argued that Washington does not force states and schools to do anything, but only attaches rules and regulations to federal money that states and districts may turn down. By essentially demanding that all states and districts adopt specific standards, however, Washington would be exceeding even the unconstitutional power it has accumulated under ESEA, violating a stipulation that has been in federal education law—including NCLB—since day one: "Nothing in this Act shall be construed to authorize an officer or employee of the Federal Government to mandate, direct, or control a State, local education agency, or school's curriculum, program of instruction, or allocation of State or local resources."⁶⁹

Measurement

There is a great deal of debate within the education community over the importance of "hard" learning, such as memorizing facts or being able to complete mathematical operations, and "soft" learning of such skills as creativity and critical thinking. Most standardized tests, by their very nature, primarily measure the former. But there is no convincing evidence that such hard learning is necessarily the most important outcome of an education system.

Given that all people are unique, the goals that they have for their education and that of their children is potentially as varied as the people themselves. Assume, though, that the

end goal for all people is happiness, whether achieved through a well-paying job, doing something that is intellectually fulfilling, or just sitting on a couch watching television all day. Taking happiness maximization as a reasonable final aim for education, it appears at least on the surface that international academic assessments are poor measures of whether or not people are getting what they want out of education. Of the 10 top-finishing countries on the 2007 TIMSS eighth-grade mathematics exam, only four—South Korea, Singapore, Japan, and the United States—ranked among the top 50 in "subjective well-being," and among those the United States ranked the highest by far on happiness.⁷⁰ Of course, a lot beyond education goes into happiness, and measuring that feeling is a tricky and imperfect undertaking, but these results at least plausibly suggest that an education system focused on achievement demonstrable by standardized test scores might not be optimal.

Coping with Diversity

A major roadblock for establishing centralized standards in any political system other than a dictatorship is getting enough people to agree on their substance. What should they include? What shouldn't they include? Who should write them? This becomes increasingly difficult the more linguistically, religiously, ideologically, and ethnically diverse a nation's population. With that in mind, it should be no surprise that many of the nations that have national standards and tend to perform best on international exams also tend to be highly homogeneous.

Consider Finland, the top-place finisher on the 2006 PISA assessment. With only about 5 million people, it has a population less than 2 percent that of the United States. In terms of mobility, it has net migration of 0.68 immigrants per 1,000 people, versus 4.31 per 1,000 Americans. It only has two major ethnic groups—Finns are roughly 93 percent of the population and ethnic Swedes about 6 percent—while the United States has a population that is predominantly "white" but 13 percent

African-American, 15 percent Hispanic, and 4 percent Asian. Religiously, about 83 percent of Finns belong to the Lutheran Church of Finland and most of the rest profess no religion. In contrast, about 51 percent of Americans are Protestant—which includes many denominations—24 percent Roman Catholic, 12 percent unaffiliated, 1.7 percent Jewish, and the rest members of several smaller religious groups. Finally, roughly 91 percent of Finns speak Finnish, 6 percent speak Swedish, and there are small Sami- and Russian-speaking minorities. In the United States, about 82 percent of people speak English, 11 percent Spanish, and the rest many other languages.⁷¹

How about South Korea, another regular top-place finisher? Its total population is 48.5 million people, about 16 percent the size of the United States. Rather than gaining migrants, South Korea is losing them. Ethnically, South Korea is almost completely homogeneous. Linguistically, everyone speaks Korean. The only way in which South Korea is remotely similar to the United States is in religious diversity, with about one-quarter of its people Christians (with significant numbers of Protestants and Roman Catholics among them), about 23 percent Buddhists, and half claiming no religious affiliation.⁷²

Finally, consider Germany, which although not a high-flyer on international assessments, is the nation that William Schmidt points to as proof that the United States can create and adopt coherent, rigorous national standards.

With about 82 million people, Germany's population is roughly a quarter the size of the United States' but bigger than South Korea's and much larger than Finland's. It has 2.19 immigrants per 1,000 people, a number about half that of the United States but much larger than Korea and Finland. Ethnically, about 92 percent of Germany's population is German, 2.4 percent Turkish, and the rest a smattering of other groups. This could be considered on par with Finland, although the gulf between ethnic Finns and Swedes is not as large as between Germans and Turks. The population of the United States, however, is much more diverse

than that of Germany. Religiously, about one-third of Germans are Protestants, one-third Roman Catholics, 3.7 percent Muslim, and the rest unaffiliated—diversity that is close to the United States and much greater than Finland. Finally, the only native language spoken in Germany is German, another significant departure from the U.S. experience but also less diverse than Finland.⁷³

Clearly, two of the top performers on international assessments are very homogeneous, and Germany—which performs roughly on par with the United States and is moving toward national standards—is more diverse than the top performers but appreciably more homogeneous than the United States. This suggests that diversity militates against establishing national standards and, perhaps, high performance on standardized tests. Because people of different backgrounds often have different priorities and demands for education, it is logical to expect that the greater the diversity of the people falling under a single schooling authority, the greater the conflict, the less coherent the curriculum, and the worse the outcomes.

Beyond just logic, while there is no systematic empirical support for the assertion above, some broad evidence is highly suggestive. Political scientist Robert Putnam has found that greater social capital—essentially, the bonds of trust between people—is correlated with better academic outcomes, and that greater social capital is found in areas with *less* diversity.⁷⁴ Similarly, sociologist James Coleman has reported that after controlling for variables such as students' socioeconomic status, Roman Catholic schools have better academic outcomes because they are “closed” systems in which all actors—students, parents, and school employees—share important norms and values.⁷⁵

What Can Be Done?

Culture, politics, diversity—all of these forces militate against the success of national standards in the United States. But if not

Culture, politics, diversity—all of these forces militate against the success of national standards.

The first argument for educational freedom is that the evidence supporting it is broad and convincing.

national standards, what is the solution to our educational woes? We have tried local control of public schools. We have tried state control. What else can be done?

The question, ultimately, is what type of reform can both successfully navigate the cultural, political, and diversity shoals on which so many public-schooling reforms have run aground and deliver optimal educational outcomes? The answer is to replace public *schooling*—in which government not only ensures that all children can access education, but also provides the schools—with true public *education*. Let parents control education funds either through vouchers or, preferably, personal and donation tax credits, and let them freely choose among autonomous schools and other educational options.⁷⁶ Let education work as a free market, in which consumers purchase services and products according to their individual needs and desires, and suppliers compete through quality, specialization, price, and innovation.

The Empirical Support for Educational Freedom

The first argument for educational freedom is that, unlike national standards, the empirical evidence supporting it is broad and convincing.

For one thing, there is a wealth of evidence showing that free markets, overall, meet the needs and desires of people much better than command economies, which is essentially what public schooling is.⁷⁷ Free markets foster competition that drives innovation and improvement, all while enabling consumers with unique needs to seek out producers that are willing and able to meet those needs.

In education, the research is also deep and convincing; including evidence based on standardized achievement tests. Ten prominent empirical studies have analyzed voucher programs using the “gold-standard” social science technique of randomly assigning subjects to “treatment” or “control” groups. Nine of those studies found that at least one subgroup of students receiving vouchers did better on academic assessments than those who applied

for but did not receive vouchers, while in no subgroup did voucher students do worse. In the tenth study, the results were a wash.⁷⁸

Comparing the outcomes of education systems around the world based on how freely consumers and providers can interact also bears this out. Reviewing 65 studies that reported over 156 separate statistical findings, Andrew Coulson found that findings favoring free-market provision of education outnumbered those favoring government-monopoly provision by a ratio of 15-to-1.⁷⁹ Notably, virtually all of these studies compared school systems in close proximity to one another *within* individual nations, which would mitigate the potentially confounding effects of cultural and other difficult-to-observe variables. Similarly, education professor James Tooley has found low-cost private (often for-profit) schools to be ubiquitous in many of the world’s poorest slums, and that they regularly outperform “free” government schools on assessments of such things as native-language mastery, English, and mathematics.⁸⁰

Going Beyond What Is Testable

So choice produces better results on achievement assessments. However, allowing people *not* to fixate exclusively on easily tested outcomes might also enable them to focus on other pursuits or skills best suited to their abilities and ultimate desires. As already discussed, Americans typically express a greater degree of happiness than their peers in nations with more centralized, test-driven education systems. If enabling people to maximize their happiness is the end goal of education—and anything else we do, for that matter—then allowing students to freely pursue their own educational interests (with the guidance of parents or other advisers) would likely produce the best educational outcomes, if not necessarily the best test scores.

Recent actions in countries that often do quite well on international examinations—nations that many national-standards advocates hope to emulate—suggest that they are, perhaps, coming to believe that education

cannot and should not be reduced simply to test scores. For instance, although Japan has recently retrenched somewhat, as early as the 1970s its education ministry started encouraging schools to put more emphasis on such soft skills as “critical thinking.” In 2002, Japan shrank its school week from six days to five, and reduced the content of its national curriculum by about 30 percent. It also introduced a class called “Integrated Studies” that features more independent, student-driven work.⁸¹

Starting in 1999, Singapore, another top performer, reduced its national curriculum by about a third, doing so in order to “provide room for teachers to implement the key initiatives announced in 1997, namely the infusing of thinking skills and integrating the use of Information Technology in lessons.”⁸² In 2001, Singapore further reformed its curriculum, focusing it even more on teaching students to “think creatively and apply knowledge innovatively.” It also introduced ability-based tracking.⁸³ Finally, funded in part by Korean education ministries, Korean teachers have been regularly coming to the United States to learn how to teach creativity and critical thinking in their schools.⁸⁴

Does this mean that easily tested subjects have been deemed irrelevant in these high-flying nations? And couldn’t administrators, teachers, and other interest groups just be trying to make their lives easier with these reforms, just as they do in the United States?

Movement away from strict standards and testing does not prove anything, of course. These nations probably still strongly support centralized standards and testing, and parents might vehemently object to softening curricula. Still, that these nations with long, proven track records of supporting tough centralized standards and testing have moved to loosen their systems must not be ignored—it could mean that they have recognized real shortcomings of their systems.

Japan, for instance, suffered the “lost decade” of the 1990s: a long period of economic malaise that many experts have at least

partially attributed to the nation’s lack of creative thinking and hidebound social structures.⁸⁵ Bolstering this is the fact that Japan has not been outpacing the United States economically, its significantly higher test scores notwithstanding. In 1980, Japanese GDP per capita, adjusted for purchasing power, was \$8,901. In the United States it was \$12,255. Since then, per capita Japanese GDP has grown about \$855 per year, versus almost \$1,180 in the United States.⁸⁶

Minimizing Poisonous Politics

In addition to enabling individuals to pursue educational options best suited to their desires and needs, free-market education avoids the political fighting that has often doomed standards-based reforms. By taking control of education away from government, much of the politics that handicaps the delivery of education would automatically be removed.

Currently, because all taxpayers have to support government schools, almost every school or district decision is politicized. District budgets constantly spark conflict. There are incessant curricular and pedagogical battles over everything from the grade at which students start to use calculators to what reading programs schools employ.⁸⁷ And of course there is the asymmetrical power problem, in which the people employed by the public schools have disproportionate power compared to those people whom the schools are supposed to serve.

Give parents control over education dollars and let them choose among autonomous schools, and these problems would diminish considerably. District budgets would be irrelevant because schools would be funded based on their ability to attract customers, not their allocations of public dollars. Curricular and pedagogical battles would peter out as schools independently chose the curricula they thought best and parents selected the schools with programs they thought most effective for their children. Finally, no longer would going through a labyrinthine and stacked political process be

Free-market education avoids the political fighting that has doomed standards-based reforms.

No one person or group has a monopoly on truth or understanding.

the only avenue through which “customers” could try to fix education problems; they would be able to execute immediate accountability by taking their children and the money to educate them out of unsatisfactory schools and putting them elsewhere.

Even with full choice, it should be noted, there would still be politics in education. For one thing, there would no doubt always be people calling for greater or less regulation of schools. In addition, what the minimum educational requirements should be for all children would likely be a contested question. Finally, depending on the mechanism used to deliver school choice, there would no doubt be recurring disputes about the maximum voucher or tax credit size, as well as eligibility criteria.

Even with those lingering problems, however, the amount of education subject to public-policy decisions would be much smaller than under the status quo. And as Andrew Coulson has empirically demonstrated, if choice is delivered through tax credits rather than vouchers—meaning no tax dollars go through government and the money heads to recipients according to the free choice of the sender—onerous regulation is much less likely.⁸⁸

Dealing with Diversity

Finally, the last great stumbling block that makes improving government schooling very difficult is diversity—ethnic, religious, and ideological. It has led to the neutering of public school curricula and the removal of potentially contentious—but also important and interesting—content from the nation’s schools. As Ravitch has documented, drives to eliminate objectionable content from curricula have come from all parts of the social and political spectra, with conservatives generally attacking morally objectionable, or insufficiently “American,” content and liberals decrying anything that could be considered critical or insensitive to various minority groups. This has resulted in textbooks that are devoid of engaging content and curricula designed to skirt controversy.⁸⁹ And this cur-

ricular denuding is not just demonstrated in overt actions. Based on widespread anecdotal evidence, it appears that to avoid conflict many teachers skip over controversial topics such as evolution even when state standards require that they be taught.⁹⁰ Perhaps that is why, despite religious instruction having long been banished from public schools, close to half of all Americans believe that human beings and other living creatures were created, most likely by God, in their present forms.⁹¹

How would school choice negotiate these treacherous shoals? In the same way it would avoid so many political fights: by letting people select educational options compatible with their diverse norms and backgrounds rather than requiring them to fight for control of a single system. Parents who want their children to learn religious explanations for human origins would be able to patronize like-minded schools, as would the strictest atheists. Hispanic families that desire instruction in Hispanic history and language could seek schools that teach them. Libertarians who want their children taught that the New Deal prolonged the Great Depression could look for schools that taught that. Then, rather than militating against any and all children hearing things that are potentially controversial, diversity would be fully accommodated and children would be able to get coherent instruction.

With Choice, Would There Be Standards?

One of the most common objections to choice—and one that is especially germane to the topic of national standards—is that free-market delivery of education would result in an absence of standards. It would create a state of relativistic, educational anarchy, according to critics.

It is true that choice would potentially let “a thousand flowers bloom” and that ideas and knowledge that might turn out to be incorrect would and could be taught. This latter possibility is something that national-standards advocates implicitly assert that they could prevent by choosing the “best” curriculum for everyone. But no one person

Conclusion

or group of people has anything close to a monopoly on truth or understanding, and the best system to deal with such imperfection is one in which power is diffuse, not concentrated. It allows ideas to compete, and those that turn out to be wrong would not have been imposed on everyone.

That said, it is inaccurate to assert that widespread standards could not and would not be adopted in a laissez-faire education system. Consider the long period of American history—from the arrival of the first British colonists to the common-schools movement in the mid-1830s—during which most education was delivered in homes, by churches, and in other nongovernmental arrangements. During that time the very “American” language that we speak was standardized by just a few commercially available tools. Noah Webster was explicit in his aims to create a standard American language and forge a unique American culture with his *American Spelling Book* and *American Dictionary of the English Language*, and he achieved great success in that endeavor. Indeed, he had sold 20 million copies of his speller by 1829, although by 1830 the total U.S. population was less than 13 million. Several other people also wrote and sold readers during this time, and the renowned and ubiquitous *McGuffey’s Reader* was published in 1836.⁹²

Today we have countless examples of nongovernmental standardization both inside and outside of education. Millions of students around the country take the SAT or ACT, as well as Advanced Placement or International Baccalaureate tests, for the most part by choice. Similarly, *U.S. News and World Report*, *Forbes*, the Princeton Review, and numerous other organizations provide guides and rankings to colleges, which unlike K–12 schools, are chosen by students. Outside of education we have Underwriters Laboratory, *Consumer Reports*, franchising, and sundry other mechanisms setting standards for innumerable items and services we use every day. Standards are ubiquitous in free markets.

The argument on behalf of national education standards is alluringly simple: set high standards that all American schools and students must meet, and all students will achieve at high levels. But actually setting high standards and getting students to meet them is extremely difficult. The opposition of the most politically powerful interests in education must be overcome. Crippling conflicts between groups with different religious, ethnic, and ideological concerns must be avoided. And finally, a culture that is averse to intense focus on academics and the kind of “hard,” largely mathematical and scientific knowledge that is easy to capture on a test must be transformed.

Much of this explains why what little comparative research there is on national standards is inconclusive. Nations vary markedly not just in terms of economics and politics, but culture and other variables that are extremely difficult, if not impossible, to cleanly control for in social research. The same applies to subnational political units, such as U.S. states and Canadian provinces. And it is hardly a settled question that test scores such as those used in standards research even encompass all the most desired educational outcomes.

For all of these reasons, the road to successful education reform appears to go in the opposite direction of greater top-down control. The key appears to be to give education funding to parents, allow schools autonomy, and as a result make schools respond to the needs and demands of parents and children. That would solve the asymmetrical power problem, forcing educators to satisfy customers rather than use politics to get their way. It would prevent paralyzing political, cultural, religious, or ethnic conflicts that force lowest-common-denominator standards. And it would lead to standards that would be meaningful, but also sufficiently flexible so that unproven ideas could compete, and inevitable human failures would not be inflicted on everyone. In other words, rather than centralizing crippling govern-

Standards are ubiquitous in free markets.

ment power even further, it would truly reform the education system.

Notes

1. U.S. Department of Education, *A Nation at Risk*, archived information, April 1983, www.ed.gov/pubs/NatAtRisk/risk.html.
2. Among several major reports calling for national standards over the last few years are Chester E. Finn Jr., Liam Julian, and Michael Petrilli, *To Dream the Impossible Dream: Four Approaches to National Standards and Tests for America's Schools* (Washington: Thomas B. Fordham Foundation, 2006); The Commission on No Child Left Behind, *Beyond NCLB: Fulfilling the Promise to Our Nation's Children* (Washington: The Aspen Institute, 2007), p. 127; and National Governors Association, Council of Chief State School Officers, and Achieve, Inc., *Benchmarking for Success: Ensuring U.S. Students Receive a World-Class Education* (Washington: National Governors Association, 2008). Among recent high-profile individual endorsers of national standards have been former U.S. secretaries of education Rod Paige and William J. Bennett, "Why We Need a National School Test," *Washington Post*, September 21, 2006; and American Federation of Teachers president Randi Weingarten, "The Case for National Standards," *Washington Post*, February 16, 2009.
3. Council of Chief State School Officers and National Governors Association, "Common Core State Standards K-12 Work and Feedback Groups Announced" press release, November 10, 2009.
4. Alyson Kline, "To Duncan, Incentives a Priority," *Education Week*, January 30, 2009, www.edweek.org/ew/articles/2009/02/04/20duncan.h28.html?r=2063662804; U.S. Department of Education, "Duncan Offers Stimulus Funds for States to Develop Rigorous Assessments Linked to Common Standards" (press release, June 15, 2009).
5. Constitutionally, it is clear that states have authority over education. Article I, Section 8, of the U.S. Constitution gives the federal government no power to govern American education, and the 10th amendment reinforces that in stating that "the powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people." In addition, even analysts who reject this interpretation of the Constitution typically acknowledge that education has traditionally been the domain of state governments, not Washington.
6. Finn et al., p. 11-12.
7. Victor Bandeira de Mello, Charles Blankenship, and Don McLaughlin, *Mapping State Proficiency Standards Onto NAEP Scales: 2005-2007* (Washington: National Center for Education Statistics, October 2009).
8. Duncan quoted in Oliver Staley and Molly Peterson, "Duncan to Spend Billions to 'Transform' U.S. Schools," *Bloomberg.com*, April 16, 2009, www.bloomberg.com/apps/news?pid=20601103&sid=aXp_0t1dgo_Q&refer=us.
9. National Governors Association, Council of Chief State School Officers, and Achieve, Inc., p. 7.
10. Robert M. Costrell, "A Simple Model of Educational Standards," *American Economic Review* 84, no. 4 (1994): 957-64.
11. Robert M. Costrell, "Can Centralized Educational Standards Raise Welfare?" *Journal of Public Economics* 65 (1997): 271-93.
12. Conference Board, Corporate Voices for Working Families, Partnership for 21st Century Skills, and Society for Human Resource Management, *Are They Really Ready to Work: Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century U.S. Workforce*, 2006, www.21stcenturyskills.org/documents/FINAL_REPORT_PDF09-29-06.pdf.
13. Andrew Flagel, "Admission to College: Does My High School Matter?," Not Your Average Admissions Blog, notjustadmissions.com/2007/11/01/admission-to-college-does-my-high-school-matter/, November 1, 2007.
14. Diane Ravitch, *National Standards in American Education: A Citizen's Guide* (Washington: Brookings Institution Press, 1995), pp. 25-27.
15. *Ibid.*, pp. xxiii-xxiv.
16. NEA and AFT headquarters employment count derived from the U.S. Department of Labor's union "Officer/Employee Search" database at kcerds.dol-esa.gov/query/getOfficerEmployeeQry.do. Searching under the name of the union and "national headquarters" generates an employment report for each "employee" or "officer" employed by the union during a given fiscal year. This likely overcounts the total number of employees at any given time during the year due to normal employee turnover, but it nonetheless provides a general sense of both unions' national presence.
17. For an excellent discussion of the history of the PTA and de facto NEA control over the group, see Charlene K. Haar, *The Politics of the PTA* (New Brunswick, NJ: Transaction Publishers, 2002), pp. 67-84.

18. LM-2s are available at the U.S. Department of Labor, "Officer/Employee Search" database, kcerds.dol-esa.gov/query/getOfficerEmployeeQry.do.
19. Ballot Initiative Strategy Center, "What We Do," www.ballot.org/pages/what_we_do.
20. Economic Policy Institute, "A Broader, Bolder Approach to Education" (report summary, June 25, 2009), www.boldapproach.org/report_20090625.html.
21. Fairtest, "The Case against High Stakes Testing," www.fairtest.org/k-12/high+stakes, and statements from unions in California, Massachusetts, Missouri, Nebraska, and Washington State at www.fairtest.org/Organizations+and+Experts+Opposed+to+High+Stakes+Testing.
22. Paul Houston, American Association of School Administrators, "High-Stake Testing Overwhelmingly Gets an 'F'" press release, June 19, 2000.
23. Vincent L. Ferrandino and Gerald N. Tirozzi, "Principals' Perspective: An Open Letter to the 110th Congress," www.principals.org/s_nasspb.in.asp?CID=1237&DID=55577&DOC=FILE. PDF.
24. National Council of Teachers of English, "Resolution on Urging Reconsideration of High Stakes Testing" (position statement), www.ncte.org/positions/statements/highstakestestrescons.
25. Chester E. Finn Jr., Liam Julian, and Michael Petrilli, *The State of State Standards 2006* (Washington: Thomas B. Fordham Foundation, 2006), p. 16.
26. On the last point, national-standards supporters will argue that polls show that Americans support standardized testing and even a national test. See, for instance, Phi Delta Kappa/Gallup, "Highlights of the 2009 Phi Delta Kappa/Gallup Poll," *Phi Delta Kappan*, September 2009, p. 8. However, while respondents might like common standards and testing in the abstract, it is entirely possible they don't like it when it becomes reality. Hence, there is high approval for "accountability," but low approval for No Child Left Behind (NCLB). For a good discussion of this, see David J. Hoff, "Polls find NCLB is Unpopular, But Accountability Isn't," *NCLB: Act II*, August 22, 2008, blogs.edweek.org/edweek/NCLB-ActII/2008/08/dissecting_polls_nclb_is_unpop.html.
27. Seymour Martin Lipset, "Still the Exceptional Nation?" *Wilson Quarterly* 24, no. 1 (Winter 2000).
28. J. M. Roberts, *A Short History of the World* (New York: Oxford University Press, 1993), pp. 196 and 206. International development scholar Francis Fukuyama notes that though the Japanese have typically been more centered on sub-state-level groups than the Chinese, they too have a very collective-oriented culture, as well as a culture of "lifetime employment" with one company that drives fixation on doing well on college entrance exams and getting into a well-regarded school. Francis Fukuyama, *Trust* (New York: Free Press, 1995), pp. 54 and 186-88.
29. Tom Loveless, *The 2001 Brown Center Report on American Education* 1, no. 2 (September 2001): 18-28, and *The 2002 Brown Center Report on American Education* 1, no. 3 (September 2002): 16-28.
30. Harold W. Stevenson, Chuansheng Chen, and Shin-Ying Lee, "Mathematics Achievement of Chinese, Japanese, and American Children," *Science* 259, no. 5091 (January 1, 1993): 53-58.
31. For an in-depth discussion of the impact of culture on academic outcomes, including, specifically, Asian students, see Abigail and Stephan Thernstrom, *No Excuses: Closing the Racial Gap in Learning* (New York: Simon and Schuster, 2003).
32. Weingarten.
33. Trends in International Mathematics and Science Study (TIMSS) mathematics scores from Institute for Education Sciences, "Trends in International Mathematics and Science Study: Table 1. Average mathematics scores of fourth- and eighth-students, by country: 2007," nces.ed.gov/timss/table07_1.asp, and "Trends in International Mathematics and Science Study: Table 3. Average science scores of fourth- and eighth-students, by country: 2007," nces.ed.gov/timss/table07_3.asp. Nations with and without national standards are identified in V. S. Mullis et al., eds., *TIMSS 2007 Encyclopedia: A Guide to Mathematics and Science Education Around the World, Volume 1 A-L*, TIMSS and Progress in International Reading Literacy Study (PIRLS) International Study Center, Exhibits 5 and 6, 2008, timss.bc.edu/TIMSS2007/encyclopedia.html. Note that the *Encyclopedia* identifies the presence of national "curriculum." For consistency, and since the term "national standard" is somewhat vague when used in policy discussions, the term "standard" has been substituted for "curriculum" in this paper.
34. Scores from TIMSS International Study Center, "TIMSS Highlights: February 1998," timss.bc.edu/timss1995i/TIMSSPDF/C_Hilite.pdf. Level of curricular decisionmaking from Ina V.S. Mullis et al., *Mathematics and Science Achievement in the Final Year of Secondary School: IEA's Third International Mathematics and Science Study* (Chestnut Hill, MA: Boston College, 1998), p. 25.
35. Scores from Organization for Economic

- Cooperation and Development, *PISA 2006 Science Competencies for Tomorrow's World*, Executive Summary, p. 22, www.oecd.org/dataoecd/15/13/39725224.pdf. Standards-based external examinations data are from *PISA 2006 Science Competencies for Tomorrow's World*, Table 5.2, www.oecd.org/dataoecd/30/63/39704360.xls.
36. *PISA 2006 Science Competencies for Tomorrow's World*, Table 5.22.
37. John H. Bishop, "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning," *International Journal of Educational Research* 23, no. 8 (1996): 653–752. In "The Effect of National Standards and Curriculum-Based Exams on Achievement," *The American Economic Review* 87, no. 2 (1997): 260, Bishop refers to "Curriculum-Based External Exit Examinations," but both that and "Curriculum-Based External Examinations" refer to standards-and-testing regimes with high-stakes consequences for secondary students.
38. Bishop, "The Impact of Curriculum-Based External Examinations," pp. 698–99.
39. *Ibid.*, p. 697.
40. *Ibid.*, pp. 697–99.
41. *Ibid.*, p. 699.
42. *Ibid.*, p. 260.
43. John H. Bishop, "Are National Exit Examinations Important for Educational Efficiency?" *Swedish Economic Policy Review* 6 (1999): 349–98.
44. *Ibid.*, pp. 374–75.
45. Ludger Woessmann, "Central Exit Exams and Student Achievement: International Evidence," in *No Child Left Behind: The Politics and Practice of School Accountability*, ed. Paul E. Peterson and Martin R. West (Washington: Brookings Institution Press, 2003), pp. 292–323. Woessmann performed a similar analysis in "Schooling Resources, Educational Institutions and Student Performance: the International Evidence," *Oxford Bulletin of Economics and Statistics* 65, no. 2 (2003). He looked at the presence of centralized exams as one among many institutional and resource variables and reached similar conclusions.
46. Woessmann, p. 304.
47. Assignment of countries to regions provided in an e-mail to the author by Ludger Woessmann.
48. Ludger Woessmann, "How Central Exams Affect Educational Achievement: International Evidence from TIMSS and TIMSS-Repeat," publication number PEPG/02-10 (Cambridge, MA: Program on Education Policy and Governance, Harvard University, 2002).
49. Hendrik Jürges and Kerstin Schneider, "International Differences in Student Achievement: An Economic Perspective," *German Economic Review* 5, no. 3 (2004): 371.
50. William H. Schmidt, Richard Houang, and Sharif Shakrani, *International Lessons about National Standards* (Washington: Thomas B. Fordham Institute, 2009), p. 29.
51. William H. Schmidt, "Panel 1: National Standards in Other Nations," (remarks at Fordham Foundation conference on national standards, May 5, 2009), www.edexcellence.net/detail/event.cfm?event_id=4&id=317.
52. Organization for Economic Cooperation and Development, "The Programme for International Student Assessment (PISA)," 2007, www.oecd.org/dataoecd/15/13/39725224.pdf.
53. In "How Well are American Students Learning?" *The 2008 Brown Center Report on American Education* (Washington: Brookings Institute, 2009), p. 16, Tom Loveless offers important insight, for instance, into the Program for International Student Assessment (PISA) science literacy test. He reports that it does not just test knowledge, but "considers students' attitudes, beliefs, and values part of literacy. Questionnaires are administered to students and some attitudinal items are interspersed in the portion of the tests assessing knowledge." This demonstrates the importance of not assuming that results on any given test reflect desired educational outcomes, and substantiates concerns that testing can become politicized—an especially dangerous problem were a single test to be imposed on the entire nation.
54. Bishop, "The Impact of Curriculum-Based External Examinations," pp. 711–24.
55. Audrey L. Amrein and David C. Berliner, "The Impact of High-Stakes Tests on Student Academic Performance: An Analysis of NAEP Results in States with High-Stakes Tests and ACT, SAT, and AP Test Results in States with High School Graduation Exams," publication number EPSL-0211-126-EPRU (Tempe, AZ: Education Policy Studies Laboratory, Arizona State University, December 2002). An earlier version of the study is "High-Stakes Testing, Uncertainty, and Student Learning," *Education Policy Analysis Archives* 10, no. 18 (March 28, 2002), epaa.asu.edu/epaa/v10no18.
56. Barak Rosenshine, "High-Stakes Testing: Another Analysis," *Education Policy Analysis Ar-*

- chives* 11, no. 24 (August 4, 2003), epaa.asu.edu/epaa/v11n24/.
57. Audrey Amrein-Beardsley and David C. Berliner, "Re-analysis of NAEP Math and Reading Scores in States with and without High-Stakes Tests: Response to Rosenshine," *Education Policy Analysis Archives* 11, no. 25 (August 4, 2003), <http://epaa.asu.edu/epaa/v11n25/>.
58. Margaret E. Raymond and Eric A. Hanushek, "High-Stakes Research," *Education Next*, Summer 2003, pp. 48–55.
59. Martin Carnoy and Susanna Loeb, "Does External Accountability Affect Student Outcomes? A Cross-State Analysis," *Educational Evaluation and Policy Analysis* 24, no. 4 (Winter 2002): 313.
60. Carnoy and Loeb, p. 320.
61. *Ibid.*, pp. 314–15.
62. Henry Braun, "Reconsidering the Impact of High-stakes Testing" *Education Policy Analysis Archives* 12, no. 1 (January 5, 2004), [epaa/v12n1/](http://epaa.asu.edu/epaa/v12n1/).
63. Hendrik Jürges, Kerstin Schneider, and Felix Büchel, "The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany," *Journal of the European Economic Association* 3, no. 5 (September 2005): 1153–54.
64. Eric A. Hanushek and Margaret E. Raymond, "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management* 24, no. 2 (2005): 302.
65. Common Core State Standards Initiative, www.corestandards.org/.
66. Diane Ravitch describes well how federal efforts to create national standards in the 1990s degenerated into nationwide acrimony in *Left Back: A Century of Battles Over School Reforms* (New York: Touchstone, 2000), pp. 429–37.
67. Kress quoted in Finn, Petrilli, and Julian, *To Dream the Impossible Dream: Four Approaches to National Standards and Tests for America's Schools*, p. 20.
68. For more on the constitutionality question, see Neal McCluskey, *Feds in the Classroom: How Big Government Corrupts, Cripples, and Compromises American Education* (Lanham, MD: Rowman and Littlefield, 2007), pp. 129–64.
69. U.S. Public Law 107-110, 107th Cong., 1st sess., January 8, 2002.
70. Subjective well-being ranking is from Will Wilkinson, "In Pursuit of Happiness Research: Is It Reliable? What Does It Imply for Policy?" *Cato Policy Analysis* no. 590 (April 11, 2007), p. 19.
71. Central Intelligence Agency, "Finland," *The World Factbook*, www.cia.gov/library/publications/the-world-factbook/geos/fi.html, and "The United States," www.cia.gov/library/publications/the-world-factbook/geos/us.html.
72. *Ibid.*, "Korea, South," *The World Factbook*, www.cia.gov/library/publications/the-world-factbook/geos/ks.html.
73. *Ibid.*, "Germany," *The World Factbook*, www.cia.gov/library/publications/the-world-factbook/geos/gm.html.
74. Findings correlating social capital with academic outcomes are from Robert D. Putnam, *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon and Schuster, 2000), pp. 299–306; and correlation of social capital with diversity is from Putnam, "E Pluribus Unum: Diversity and Community in the Twenty-first Century," *Scandinavian Political Studies* 30, no.2 (June 15, 2007): 149.
75. James S. Coleman, "The Creation and Destruction of Social Capital: Implications for the Law," *Notre Dame Journal of Law, Ethics, and Public Policy* 375 (1987–1988): 375–404.
76. For an excellent explanation of how and why education tax credits work, see Adam Schaeffer, "The Public Education Tax Credit," *Cato Policy Analysis* no. 605 (December 5, 2007).
77. For a powerful discussion of the overwhelming evidence of this, see Johan Norberg, *In Defense of Global Capitalism* (Washington: Cato Institute, 2003).
78. For a review of those studies with links to attain them, see Jay P. Greene, "Voucher Effects on Participants," Jay P. Greene's Blog, jaypgreene.com/2008/08/21/voucher-effects-on-participants/.
79. Andrew J. Coulson, "Comparing Public, Private, and Market Schools: The International Evidence," *Journal of School Choice* 3 (2009): 31–54.
80. James Tooley, *The Beautiful Tree: A Personal Journey Into How the World's Poorest People Are Educating Themselves* (Washington: Cato Institute, 2009), pp. 178–82.
81. Christopher Bjork and Ryoko Tsuneyoshi, "Education Reform in Japan: Competing Visions for the Future," *Phi Delta Kappan* 86, no. 8 (April 2005): 620–21.

82. Singapore Ministry of Education, "Content Reduction in the Curriculum" press release, July 16, 1998, www.moe.gov.sg/media/press/1998/980716.htm.
83. Teo Chee Hean (speech on national curriculum changes, October 8, 1999), <http://www.moe.gov.sg/media/speeches/1999/sp081099.htm>.
84. Colorado State University Department of Public Relations, "Colorado State University Welcomes 40 Korean Teachers for One Month of Innovative Science Education Training" (press release, www.news.colostate.edu/Release/557, July 15, 2005); and Indiana University, "Korean Teachers Taking in Differences, Making Notes During Month-long Visit" (press release, newsinfo.iu.edu/news/page/normal/8580.html, July 30, 2008).
85. See, for instance, Masaru Tamamoto, "Will Japan Ever Grow Up?" *Far Eastern Economic Review*, (July 10, 2009), www.feer.com/essays/2009/july/will-japan-ever-grow-up; and Kathleen Kennedy Manzo, "Trends in Japan: Japan Continues Search for Academic Triumph," *Education Week*, April 23, 2008.
86. GDP per capita adjusted for purchasing-power-parity using International Monetary Fund, World Economic Outlook Database, April 2009 edition, www.imf.org/external/pubs/ft/weo/2009/01/weo_data/index.aspx.
87. For instance, see Katherine Long, "Which Math Book To Use? A Passionate Debate Rages," *Seattle Times*, August 16, 2009, and Matthew E. Milliken, "Parents Blast 'Reading Street,'" *Herald-Sun*, September 28, 2009.
88. Andrew J. Coulson, "On the Way to School: Why and How to Make a Market in Education" (book chapter prepared for The Clemson University Conference on School Choice Program Design, forthcoming).
89. Diane Ravitch, *The Language Police: How Pressure Groups Restrict What Students Learn* (New York: Alfred A. Knopf, 2003).
90. Cornelia Dean, "Evolution Takes a Back Seat in U.S. Classes," *New York Times*, February 1, 2005.
91. Scott Keeter and Juliana Horowitz, "On Darwin's 200th Birthday, Americans Still Divided about Evolution," Pew Research Center, February 5, 2009.
92. Wayne J. Urban and Jennings L. Wagoner Jr. *American Education: A History*, 3rd ed. (Boston: McGraw Hill, 2004), pp. 79–82.

RELEVANT STUDIES IN THE POLICY ANALYSIS SERIES

- 641. **The Poverty of Preschool Promises: Saving Children and Money with the Early Education Tax Credit** by Adam B. Schaeffer (August 3, 2009)
- 629. **Unbearable Burden? Living and Paying Student Loans as a First-Year Teacher** by Neal McCluskey (December 15, 2008)
- 620. **Markets vs. Monopolies in Education: A Global Review of the Evidence** by Andrew J. Coulson (September 10, 2008)
- 618. **The Fiscal Impact of a Large-Scale Education Tax Credit Program** by Andrew J. Coulson with a Technical Appendix by Anca M. Cotet (July 1, 2008)
- 616. **Dismal Science: The Shortcomings of U.S. School Choice Research and How to Address Them** by John Merrifield (April 16, 2008)
- 605. **The Public Education Tax Credit** by Adam B. Schaeffer (December 5, 2007)
- 599. **End It, Don't Mend It: What to Do with No Child Left Behind** by Neal McCluskey and Andrew J. Coulson (September 5, 2007)

STUDIES IN THE POLICY ANALYSIS SERIES

- 660. **Lawless Policy: TARP as Congressional Failure** by John Samples (February 4, 2010)
- 659. **Globalization: Curse or Cure? Policies to Harness Global Economic Integration to Solve Our Economic Challenge** by Jagadeesh Gokhale (February 1, 2010)
- 658. **The Libertarian Vote in the Age of Obama** by David Kirby and David Boaz (January 21, 2010)
- 657. **The Massachusetts Health Plan: Much Pain, Little Gain** by Aaron Yelowitz and Michael F. Cannon (January 20, 2010)
- 656. **Obama's Prescription for Low-Wage Workers High Implicit Taxes, Higher Premiums** by Michael F. Cannon (January 13, 2010)
- 655. **Three Decades of Politics and Failed Policies at HUD** by Tad DeHaven (November 23, 2009)

654. **Bending the Productivity Curve: Why America Leads the World in Medical Innovation** by Glen Whitman and Raymond Raad (November 18, 2009)
653. **The Myth of the Compact City: Why Compact Development Is Not the Way to Reduce Carbon Dioxide Emissions** by Randal O'Toole (November 18, 2009)
652. **Attack of the Utility Monsters: The New Threats to Free Speech** by Jason Kuznicki (November 16, 2009)
651. **Fairness 2.0: Media Content Regulation in the 21st Century** by Robert Corn-Revere (November 10, 2009)
650. **Yes, Mr President: A Free Market Can Fix Health Care** by Michael F. Cannon (October 21, 2009)
649. **Somalia, Redux: A More Hands-Off Approach** by David Axe (October 12, 2009)
648. **Would a Stricter Fed Policy and Financial Regulation Have Averted the Financial Crisis?** by Jagadeesh Gokhale and Peter Van Doren (October 8, 2009)
647. **Why Sustainability Standards for Biofuel Production Make Little Economic Sense** by Harry de Gorter and David R. Just (October 7, 2009)
646. **How Urban Planners Caused the Housing Bubble** by Randal O'Toole (October 1, 2009)
645. **Vallejo Con Dios: Why Public Sector Unionism Is a Bad Deal for Taxpayers and Representative Government** by Don Bellante, David Denholm, and Ivan Osorio (September 28, 2009)
644. **Getting What You Paid For—Paying For What You Get: Proposals for the Next Transportation Reauthorization** by Randal O'Toole (September 15, 2009)
643. **Halfway to Where? Answering the Key Questions of Health Care Reform** by Michael Tanner (September 9, 2009)
642. **Fannie Med? Why a “Public Option” Is Hazardous to Your Health** by Michael F. Cannon (July 27, 2009)