What the Federal Government Can Do to Improve High School Performance

Russell W. Rumberger*

University of California, Santa Barbara

Paper commissioned by the Center on Education Policy, Washington, D.C., for its Project on Rethinking the Federal Role in Education

July 2009

The views expressed in this paper are those of the author.

*I want to thank Jack Jennings, Lorraine McDonnell, Jim Connell, Susan Rotermund, and five anonymous reviewers for their comments on an earlier version of this paper, as well as Beverly Bavaro and Rebecca Barns for their editorial assistance.

Executive Summary

High schools play a crucial role in preparing students for college, work, and citizenship. Yet, by many accounts, U.S. high schools are not performing any of these tasks well. This situation has prompted calls for improving high school performance. This report reviews past efforts to reform high schools, examines why those efforts have largely been unsuccessful, and suggests what the federal government can do to improve high school performance.

In order to improve the performance of U.S. high schools, it is first necessary to identify the purposes and goals of high schools and then develop suitable measures of school performance to determine the extent to which those goals are met. Only then can any serious effort be made to improve high school performance. In the current era of standards-based accountability, reform efforts have focused on raising student *academic* performance as measured by course credits, test scores, and educational credentials. Yet research studies and surveys of employers suggest students need a wide variety of non-academic as well as academic skills to be successful in college, the workplace, and in their adult lives.

A number of approaches have been developed for improving high schools, including targeted approaches that focus on specific facets of the school (instruction, student support, school restructuring); comprehensive strategies that redesign all aspects of the school or create new schools; collaborative approaches that create partnerships between schools and outside agencies; and systemic approaches that alter requirements for all schools in the system. Although the research evidence on the effectiveness of specific approaches is limited, it does suggest that no one strategy is inherently more effective than the others.

Numerous large-scale initiatives to improve the performance of high schools in the U.S. have been undertaken in the past 20 years by government agencies, foundations, non-profit organizations, and independent developers. For the most part these efforts have been unsuccessful, although there was widespread variability in both the implementation and impact of the initiatives across schools, districts, and states. Evaluations of these efforts have identified a number of factors that limited their implementation and impact, with the most important being the lack of *will* and *capacity* of both individual educators and institutions to engage in sustained improvement efforts. One implication is that strategies for improving high schools *will not be successful until critical aspects of capacity and context are improved*.

The federal government can play an important role in improving U.S. high schools by shifting its focus from short-term accountability to long-term capacity building. Specifically, the federal government should:

1. Support the development of broader indicators of student progress and outcomes, and include these indicators in the National Assessment of Educational Progress.
2. Help build the capacity of state governments and technical-assistance providers to support improvement efforts and capacity building in districts and schools.
3. Develop guidelines to insure that states do a better job of matching reform strategies to the capacity of schools and districts in need of improvement.
4. Improve coherence among federal policy initiatives, between federal and state initiatives, and between government and foundation initiatives.
5. Support the development of more comprehensive state and local data systems that not only measure educational inputs and outputs, but also district and school readiness and capacity to initiate reform as well as progress toward improving student outcomes.

High schools play a crucial and complex role in the U.S. educational system. Because high schools are the final level of compulsory schooling, they must prepare students for different destinations. For students who plan to attend college, high schools should provide the skills and knowledge to help them successfully complete more advanced credentials in the higher education system. For students who do not choose to or cannot afford to attend college, at least not immediately after high school, high schools should provide the knowledge and skills to adequately prepare them to enter the workforce. At the same time, high schools should prepare all students for a successful transition to adulthood, which includes becoming responsible, productive citizens.

By many accounts, U.S. high schools are not performing any of these tasks well. First, many students do not even graduate from high school. The federal government estimates that only 75 percent of ninth grade students enrolled in U.S. public schools in the fall of 2001 graduated four years later (Snyder, Dillow, & Hoffman, 2009, Table 106). Second, the academic performance of students who stay in school is generally low. Based on the 2007 National Assessment of Educational Progress (NAEP), only 40 percent of twelfth grade students were proficient in reading, and only 25 percent were proficient in math (Grigg, Donahue, & Dion, 2007). Third, a majority of high school students are not academically prepared for college. One study estimated that only 38 percent of high school students in the graduating class of 2002 were academically prepared for a four-year college (Greene & Winters, 2005). Another study estimated that just 25 percent of California high school students in the graduating class of 2007 had successfully completed all the courses required for admission to the stateôs four-year universities (Rogers et al., 2009). Finally, recent surveys of public high school graduates, college

instructors, and employers all found significant numbers of high school graduates underprepared for work and college (Peter D. Hart Research Associates/Public Opinion Strategies, 2005).

The poor performance of U.S. high schools threatens the countryôs international competitiveness. In his best-selling book, *The World is Flat*, The *New York Times* columnist Thomas Friedman argues that new developments in telecommunications have ñflattenedò the world, allowing entrepreneurs from large, growing economies like India and China to compete with the U.S. for a variety of highly skilled jobs in telemarketing, accounting, computer programming, and engineering (Friedman, 2005).

Increasingly, other countries that have greatly expanded their education systems are challenging the historical comparative advantage of the United States. According to data collected by the Organisation for Economic Co-Operation and Development (OECD), the United States ranked 20th among OECD countries in the percentage of students who graduated from high school in 2006, with a graduation rate of 77 percent (Organization for Economic Co-operation and Development, 2008, Chart A2.1). More disturbing, over the ten-year period from 1995 to 2005, the average graduation rate in OECD countries improved by six percentage points, from 77 to 83 percent, while the average graduation rate in the U.S. improved by a mere three percentage points, from 74 to 77 percent (Organization for Economic Co-operation and Development, 2008, Table A2.1). In addition, the U.S. ranks low in international comparisons of student achievement. For example, the U.S. ranked 25th out of 30 countries in a 2006 international assessment of mathematics literacy (Snyder et al., 2009, Table 403). Together, these data show that the U.S. is falling behind other countries in high school achievement and graduation levels, thus hampering its ability to improve training and college participation beyond high school.

In a speech to the a joint session of Congress on February 24, 2009, the President of the United States, Barack Obama, stressed the urgency for improving Americaôs high schools and reducing the nationôs dropout rate:

> In a global economy where the most valuable skill you can sell is your knowledge, a good education is no longer just a pathway to opportunity ï it is a pre-requisite. Right now, three-quarters of the fastest-growing occupations require more than a high school diploma. And yet, just over half of our citizens have that level of education. We have one of the highest high school dropout rates of any industrialized nation. And half of the students who begin college never finish. This is a prescription for economic decline, because we know the countries that out-teach us today will out-compete us tomorrowé Americaôs competitiveness demands a focus on the needs of our lowest-performing students and schools. Our middle- and high- schools must identify students at-risk of dropping out, and we must scale-up models that keep students on a path toward graduation. Reform in Americaôs lowest-performing schools must be systemic and transformational (U.S. President, 2009).

Such a call for reform is not new. There is a long history of efforts in the United States to improve the nationôs high schools. The federal government, state governments, major foundations, and independent reformers have all undertaken numerous efforts to design, fund, and implement both targeted programs for improving various aspects of high school performance and comprehensive school reform models to transform entire high schools. Yet, considering the size and scope of some of these efforts, there is little scientific evidence on either the effectiveness of high school improvement strategies or the impact of large-scale initiatives to promote high school reform. For example, in a review of 10 comprehensive school reform models in which at least half of the reform focused on secondary schools, outside evaluators found only two programs that showed significant improvements in student outcomes (Borman, Hewes, Overman, & Brown, 2003).

The federal governmentôs What Works Clearinghouse (WWC) identified 16 dropout intervention studies that met its criteria for rigorous experimental evidence, yet found only four programs that were effective in improving high school completion rates, and none of these four

programs were effective in helping students earn a regular high school diploma (U.S. Department of Education, Institute of Education Sciences, 2008). Federal and state governments, along with numerous foundations, have invested millions of dollars in promoting high school reform, but none of those efforts have made widespread improvements in student outcomes.

Why is improving high school performance so difficult? Does the difficulty lie in developing effective reform strategies, or in replicating or scaling up effective strategies? In what ways are high school improvement strategies similar or dissimilar to strategies for improving elementary and middle schools? And what role can the federal government play in promoting effective high school reform in order to improve student achievement, high school graduation rates, and better preparation for college and work?

This paper addresses these questions. First, the paper examines the goals of high schools and alternative ways to measure high school performance. Second, it reviews the various approaches for improving high school performance and the scientific evidence on their effectiveness. Third, it reviews past efforts by the federal and state governments, foundations, and private groups to promote widespread high school reform and what has been learned from those efforts. Fourth, it examines the role of policy in promoting school improvement and the shortcomings of past policy approaches. Finally, it suggests a more effective role for federal policy. Although this paper focuses on strategies for improving high schools directly, an increasing body of evidence finds that it is student performance in middle school and even elementary school that predicts high school performance (ACT, 2008; Alexander, Entwisle, & Kabbini, 2001; Balfanz, Herzog, & Mac Iver, 2007), which suggests that efforts to improve high school performance should include interventions in earlier grades. We return to this point in the concluding section of the paper.

The overall conclusion of this paper is that past efforts to reform schools, including high schools, have relied too much on accountability systems developed by the federal and state governments to pressure schools to improve test scores and not enough on building the capacity of schools to support long-term improvement in a broader and more meaningful array of educational outcomes. The paper recommends a number of actions the federal government can take to build the capacity of schools, districts, and state education agencies to support improvement, including support for more robust data systems to better measure improvements in school capacity and a broader set of performance indicators for high schools.

Measuring High School Performance

In order to improve the performance of U.S. high schools it is first necessary to identify the purposes and goals of high schools and then develop suitable measures of school performance to determine the extent to which those goals are met. Only then can any serious effort be made to improve high school performance.

Yet this issue is generally given little attention in discussions of school reform. In the current era of standards-based accountability, the focus of most reform efforts has been to raise student *academic* performance as measured by course credits, test scores, and educational credentials. But schools have always had much broader goals, and the particular goals of students, parents, teachers, administrators, policymakers, and the public at large may vary.

The Conflicting Goals of High School

There has been a longstanding debate over the purposes and goals of high schools in the United States from their inception in the 19th century. The debate concerns who should attend high

schools (whether access and attendance should be universal or selective), and a related issue of what is studied (whether the curriculum should be the same for everyone or differentiated based on the background, perceived ability, and likely destination of the students).

When high schools were first established in the late 1800s, they were selective, catering to the small portion of students who were preparing to attend college, and offered a common curriculum of academic subjects (Dorn, 1996). As enrollment surged during the two decades surrounding the turn of the 20th century, educators, business leaders, and government officials argued for a more inclusive role for high schools, one with a broader mission that emphasized practical curriculað health, ñcommand of fundamental processes,ò ñworthy home-membership,ò vocation, citizenship, leisure, and ethicsð as well as different curricula for different students.[1]

The differentiation of the high school curriculum was guided by the idea of ñsocial efficiency,ò which suggested that the role of schools was to prepare different students for different positions that they would assume in adult life. This idea was championed by what historian David Tyack labeled as the ñadministrative progressivesò and was based on what they found in influential training programs for school superintendents in such renowned universities as Columbia, Chicago, and Stanford (Tyack, 1974). The idea was greatly bolstered by the growth of group intelligence tests after World War I. These views persisted well into mid-century. In a widely cited book, *The American High School Today*, published in 1959, James Bryant Conant, former president of Harvard University, argued for a comprehensive high school that all could attend, but one that provided a differentiated curriculum that prepared students for different positionsð some for advanced schooling and some for the workplace.

[1] These goals were outlined in the 1918 report, *The Cardinal Principles of Secondary Education*, commissioned by the federal government (Dorn, 1996, p. 42).

The broad goals for high schools were mirrored in the goals for public education more generally. In the early 1950s, President Dwight Eisenhower helped secure funds from Congress for states to host conferences on the goals of education. These culminated in a 1955 national meeting where civil and business leaders together with educators recommended a list of 15 goals for public education that included:

- A general education, with increased emphasis on the physical and social sciences
- Patriotism and good citizenship
- Moral, ethical, and spiritual values
- Vocational education tailored to the abilities of each pupil and to the needs of the community and nation
- Domestic skills
- Health services for all children, including physical and dental inspections, and instruction aimed at bettering health knowledge and habits (Rothstein, Jacobsen, & Wilder, 2008, p. 27)

Support for a broad range of educational goals continues to this day. A recent survey of the general public, elected officials, and state legislators found widespread support for a set of broad educational goals (**Table 1**). Although the highest rated goal among all three groups of respondents was basic academic skills, it still garnered no more than one-quarter of the votes. Overall, there was relative agreement among all three groups of respondents that public education should meet a wide array of goals.

Further support for the value of broad educational goals comes from academic research studies and surveys of employers that suggest students need a wide variety of skills to be successful in college and in the workplace as well as to successfully transition to adulthood (Lippman, Atienza, Rivers, & Keith, 2008; Wagner, 2008; Olson, 2007). These skills include traditional academic skills, but also applied, vocational skills, as well as so-called ñsoftò skills or

ñnon-cognitiveò skills, such as punctuality, perseverance, and the social skills needed to work in groups (Heckman & Rubinstein, 2001). In fact, one recent study found that improvements in a range of non-academic skills were more valuable than improvements in math achievement for increasing studentsô chances for enrolling in and completing postsecondary programs, and for increasing earnings eight years after high school (Deke & Hiamson, 2004). A comprehensive review of the literature on healthy youth development suggests that students need an even wider range of competencies in five domainsð physical, psychological, social, cognitive, and spiritualð to ensure success in college, work, and the transition to adulthood, although there is agreement that only some competencies, such as high expectations and good communication skills, are necessary in all three areas (Lippman et al., 2008).

Despite widespread support for broad educational goals, the requirements for high school graduation and accountability systems established by the federal and state governments to improve school performance focus almost exclusively on academic outcomes.

Most states specify both a minimum number and the specific set of courses required for students to earn a high school diploma, although school districts can add additional requirements. These requirements vary among states and over time. In 2006, among states that specified course requirements, the number of course credits varied from a low of 13 in such states as California, Wisconsin, and Wyoming to a high of 24 in such states as Alabama, Florida, and South Carolina (Snyder et al., 2009, Table 167). Alabama appears to have the most rigorous requirements: four yearlong courses in English, social studies, science, and mathematics, together with eight additional courses. This translates into four academic and two additional courses for each of the four years of high school. Such requirements leave little room for error; students who fail to earn

six credits per year would not progress in school and would run the risk of not graduating in the expected four years.

There is a growing movement in many states to not only specify the number and types of courses needed for high school graduation, but also the content standards and sequence of courses needed to become ñcollege and workplace ready.ò This effort is being led by the American Diploma Project (ADP), a joint initiative of three prominent national organizationsð Achieve, The Education Trust, and The Thomas B. Fordham Foundation. In its report, *Ready or Not: Creating a High School Diploma that Counts* (2004), the Project argues that all students, no matter what their future destination, should take the same academic curriculum in high school:

> Successful preparation for both postsecondary education and employment requires learning the same rigorous English and mathematics content and skills. No longer do students planning to go to work after high school need a different and less rigorous curriculum than those planning to go to college. In fact, nearly all students will require some postsecondary education, including on-the-job training, after completing high school. Therefore, a college and workplace readiness curriculum should be a graduation requirement, not an option, for all high school students (pp. 8-9).

The report goes on to suggest that states should support and encourage different approaches or ñmultiple pathwaysò to help students meet these standards, including vocational programs, project-based learning, charter schools, and advanced coursework, such as Advanced Placement and International Baccalaureate Programs (p. 10). Some states have gone further, mandating a broader curriculum for all students that includes career and technical education (American Diploma Project, 2007).

In addition to earning the required number and type of course credits, students in many states must also pass a high school exit exam. In the 1970s, several states began using minimum competency tests, typically in a multiple-choice format, to assure that students had mastered basic skills before receiving a high school diploma. With the growth of the accountability

movement and standards-based reform in the 1990s, assessments were used to ensure that students had mastered more rigorous, grade-level content standards. In 1979, New York became the first state to require an examination before a student could receive a diploma (Warren, Jenkins, & Kulick, 2006). In the 1980s, 12 additional states added exam requirements, and four more added exam requirements in the 1990s. By 2006, 22 states required that students pass some sort of exit exam before being awarded a diploma (Snyder et al., 2009, Table 167). All of the exams are limited to core academic subjects, such as English and math, with some states also including science and history.

Some states also use exams to award higher-level diplomas. New York, for example, awards three levels of high school diplomas based on the types of course credits earned (all requiring 22 yearlong credits) and on the number of subjects and scores on the Regents Examð from a minimum score of 55 in five subjects for a Local Diploma to a minimum score of 65 in eight subjects for a Regents Diploma with advanced designation (City University of New York, 2009).

The minimum graduation requirements based on coursework and exams set by states and local districts dictate the level of performance that all students must meet in order to earn a high school diploma. Those requirements have increased over the last couple of decades, and many states have established higher requirements for future graduating classes (American Diploma Project, 2007). High school graduation requirements provide an answer to a fundamental question: What knowledge and skills should *all* high school graduates possess in order to receive a diploma?

The American Diploma Project, and the 34 states that support it, argues that all students need rigorous levels of English and math because such courses are ñprerequisites for success in

college and well-paying jobsò (American Diploma Project, 2007, p. 2). Yet the English and math skills for some of the jobs that are profiled do not require such rigorous skills. For example, the job of Events Manager requires a number of English skills, but no rigorous math skills are identified (American Diploma Project, 2004, p. 82). If such jobs do not in fact require such rigorous levels of academic coursework, why require all students to meet those levels of coursework to receive a diploma? Similarly, the suggested math curriculum in the American Diploma Projectð four years of math with coursework beyond the level of Algebra IIð exceeds the math requirement for entrance to the University of California.[2] Should all students meet this requirement before they are awarded a high school diploma, even if they do not wish to attend a four-year college? Although students who complete more rigorous coursework may have access to better jobs and be more likely to successfully complete postsecondary degrees, again the fundamental question is, what level of performance should *all* students be required to meet in order to graduate from high school? We revisit this topic later in the paper.

Federal and state accountability systems also focus exclusively on academic performance. Yet as Richard Rothstein and his colleagues point out, that was not always the case. The early versions of the federal governmentôs accountability system, the National Assessment of Educational Progress, were designed to measure a broad array of behaviors and abilities, not simply academic skills, reflecting the view of the committee and especially its chair, Ralph Tyler, that NAEP should assess any goal area to which schools devote ñ15-20% of their timeé the less tangible areas, as well as the customary areas, in a fashion the public can grasp and understandò (Rothstein et al., 2008, p. 101). But because of political pressure and budget constraints, soon after its first administration in 1969, NAEP focused almost exclusively on academic skills in reading, math, science, and history as assessed with paper and pencil tests.

[2] See: http://www.ucop.edu/a-gGuide/ag/a-g/math_reqs.html

There were repeated criticisms of this narrow focus. The National Academy of Education established a committee in 1987 to evaluate NAEP because of a concern that it was too narrowly focused on literacy and math skills that ñcould have a distorted impact on our schoolsò (Rothstein et al., 2008, p. 31). The report identified a fundamental contradiction between the goals of public education and the accountability system designed to assess it:

> At root here is a fundamental dilemma. Those personal qualities that we hold dearð resilience and courage in the face of stress, a sense of craft in our work, a commitment to justice and caring in our social relationships, a dedication to advancing the public good in our communal lifeð are exceedingly difficult to assess. And so, unfortunately, we are apt to measure what we can, and eventually come to value what is measured over what is left unmeasured. The shift is subtle, and occurs gradually. It first invades our language and then slowly begins to dominate our thinking. It is all around us, and we too are a part of it. In neither academic nor popular discourse about schools does one find nowadays much reference to the important qualities noted above. The language of academic achievement tests has become the primary rhetoric of schooling (Rothstein et al., 2008, p. 31).

Alternative Measures of High School Performance

In order to improve high schools, it is first necessary to judge how well they are currently performing. In other words, how should high school performance be measured? As pointed out in the National Academy Report, there is a dilemma between what high schools are trying to do and should be doing to better prepare young people for their future lives, versus what can and is being measured to judge how well they are doing these tasks. This section briefly describes some of the aspects of high school performance that are currently being measured and how well high schools are performing based on these measures.

It is also important to make a distinction between the performance of high school students and the performance of high schools themselves. Student performance in school is a function of both a studentsô background and the schools that they attend. As famed sociologist

James Coleman discovered in his 1966 landmark study, *Equality of Educational Opportunity*: ñfamily background differences account for much more variation than do school differencesò in student achievement (1966, reprinted in Coleman, 1990, p. 124). More specifically, Coleman found that schools only accounted for 5 percent to 38 percent of the variation in achievement among different grade levels, ethnic groups, and regions of the country (Coleman, 1990, p. 77). And although Colemanôs analysis was subjected to considerable scrutiny and debate, more recent research studies using sophisticated statistical techniques continue to find that no more than a third of the variation in student achievement lies between different schools (Raudenbush & Bryk, 2002; Rumberger & Palardy, 2004). This finding has important implications for policies designed to improve student achievement that focus on schools as the sole arena for change, a topic revisited below.

Student Performance

There are a number of current indicators that can be used to judge how well students are performing in high school. Some represent *outcome* indicators because they can be used to judge how well students meet or exceed the requirements for high school graduation; others represent *progress* indicators because they can be used to judge how well students are progressing toward the desired outcomes of high school.

Three outcome indicators are: (1) high school graduation, (2) coursework, and (3) academic achievement.

1. *High school graduation*. The first and perhaps most important outcome indicator of student performance in high school is graduation, which is defined as receiving a regular high school diploma. Students in many states can earn a high school equivalency through either a

state or national examination, with the most common national examination being the General Educational Development (GED) test administered by the American Council of Education (General Education Development Testing Service, 2008). Some states award regular diplomas, while other states award ñequivalencyò diplomas or certificates based on either the GED or state-designed examinations (see GED, Appendix A). The distinction between completing high school by earning a regular diploma and completing high school by earning an equivalency diploma is important: the economic benefits are not equivalent (Cameron & Heckman, 1993; Tyler, 2003). As a result, both the federal and state accountability systems only acknowledge students who earn regular high school diplomas.

Although defining a high school graduate is straightforward, measuring high school graduation rates is not. The nationôs governors and the federal government have all endorsed a formula known as the *four-year adjusted cohort graduation rate*, which measures the percentage of entering (first-time) ninth graders who graduate with a regular diploma in four years, adjusted for transfers in and out, émigrés, and deceased students (U.S. Department of Education, 2008). But to accurately measure such a rate requires a longitudinal student data system that can track students from ninth grade through high school, which most states do not yet have.[3] Consequently, states currently use cross-sectional data to compute graduation rates using different formulas that yield vastly different rates (Miao & Haney, 2004; Warren, 2005). In addition, the federal government uses one formula to estimate four-year graduation rates for states and districts, while the nationôs leading education newspaper, *Education Week*, uses another formula to estimate rates for states and districts.

[3] The Data Quality Campaign is a national effort to promote the development and use of state longitudinal student data systems. See: http://www.dataqualitycampaign.org/

Nobel laureate economist James Heckman examined the various sources of data used to calculate dropout and gradation rates and, after correcting for errors in previous calculations, concluded that:

- The high school graduation rate is lower than the federal government reports;
- It has been declining for the past 40 years;
- Disparities in graduation rates among racial and ethnic minorities have not improved over the last 35 years (Heckman & LaFontaine, 2008).

Graduation rates from public high schools, as reported by the National Center for Education Statistics, show similar trends. The graduation rate (using the Averaged Freshman Graduation Rate) was at its highest level in 1970 at 78.7 percent and then trended downward, reaching its lowest level of 71.0 percent in 1996 (see **Figure 1**). The rate has trended upward since that time, reaching an estimated 74.8 percent in 2009. Enrollment rates show different trends, reaching a low of 89.8 percent in 1980 and peaking at 95.2 percent in 2006. Thus, the declining rates of high school graduation are not a reflection of declining enrollment; in fact, enrollment has been increasing while graduation rates have been decreasing. What accounts for these divergent trends is not clear.

2. *Coursework*. Another useful outcome indicator of student performance in high school is measured by the number and types of courses that students complete. Although students must complete a prescribed number and types of courses to earn a high school diploma, they can and often do take more than the minimum required for high school graduation, in part to improve their prospects for entering and completing college.

A recent federal study documented trends in the number and types of course credits that graduating students earned between 1982 and 2004 (Planty, Provasnik, & Daniel, 2007). The

study found that in 2004, high school graduates earned an average of 25.8 credits, compared to 21.7 credits for graduates in 1982 (p. 7). The number of credits earned in specific subject areas also increased: 4.3 vs. 4.0 credits in English; 3.6 vs. 2.7 credits in math, and 3.2 vs. 2.2 credits in science. An increasing percentage of students completed advanced courses as well. For example, the percentage of students who completed at least one course more challenging than Algebra II increased from 26 percent in 1982 to 50 percent in 2004, and the percentage of students who completed at least one course more challenging than general biology increased from 35 percent in 1982 to 68 percent in 2004 (p. 10). In addition, the number of students taking college-level Advanced Placement courses more than doubled between 1997 and 2005 (p. 14). Another study comparing the high school transcripts of high school graduates in 1990 and 2005 confirms these trends, and further notes increases in grades and in the percentage of students who completed various curriculum levelsð specific combinations of academic coursework in English, social studies, mathematics, science, and foreign language (Shettle et al., 2007).

3. ***Academic Achievement***. A third outcome indicator is academic achievement, as measured by performance on standardized tests. Student achievement indicators provide information on what students actually know, at least as measured by the scope and validity of the tests that are employed (National Research Council, Committee on Appropriate Test Use, 1999). Most existing tests focus on academic achievement in the primary subject areas of English, math, science, and history.

Longitudinal data from the National Assessment of Educational Progress show that the average reading score for 17-year-olds in the U.S. improved slightly from 1971 to 1992, but has since declined to the same level as in 1971.[4] Average math scores for 17-year-olds declined

[4] These data are taken from NAEP website, Long-Trends Major Results. Retrieved April 8, 2009, from: http://nces.ed.gov/nationsreportcard/ltt/results2004/

slightly from 1973 to 1982, and have remained in the same statistical range since 1990. Put differently, even though the amount and rigor of high school studentsô coursework has increased substantially over the last decade and a half, test scores have remained virtually unchanged.

In addition to student outcome indicators, student performance in high school can be assessed with a number of indicators that measure studentsô progress toward graduation.

1. *Dropout Rates*. Dropping out is a major deterrent to finishing high school. Although students who drop out of school can re-enroll and earn a high school diploma, follow-up studies of dropouts find that only a fraction tend to do so. A somewhat higher proportion earn alternative credentials such as a GED (Rumberger & Rotermund, 2008).

As with graduation rates, there are different sources of data and formulas for calculating dropout rates. The U.S. Census estimates the percentage of 15- to 24-year-olds who drop out of grades 10-12 each year; this rate declined from 6.1 percent in 1972 to 3.8 percent in 2006 (Laird, Cataldi, KewalRamani, & Chapman, 2008, Table 4). Based on data collected from state education agencies, the U.S. Department of Education estimates that 3.9 percent of public school students enrolled in grades 9-12 dropped out in 2005-06 (Snyder et al., 2009, Table 107).

2. *Retention*. Another strong predictor of graduation is retention. Students who are retained are more likely to drop out and less likely to graduate than students who are not retained (Rumberger & Lim, 2008). In high school, retention occurs when students do not earn enough credits toward graduation to advance to the next grade level because they fail one or more classes. Ninth grade course failure is one of the most powerful indicators of whether students are likely to drop out (Allensworth & Easton, 2005).

Although retention data are not collected at the national level, one way to estimate the percentage of retained ninth-graders is to compare ninth grade enrollment with eighth grade

enrollment a year earlier. In 2000, there were 13 percent more ninth-graders than there were eighth-graders one year earlier (Heckman & LaFontaine, 2008, p. 16). The bias is even greater for minority students, where ninth grade enrollment was found to be 20-26 percent higher. States with longitudinal data systems also show high retention rates in ninth grade. In Texas, for example, 15.4 percent of 2006-07 ninth-graders were retained, compared to only 1.5 percent of eighth-grade students, with ninth grade retention rates at 20.8 percent for Hispanic students and 19.2 percent for African American students (Texas Education Agency, 2008, Tables 2 and 5).

School Performance

If improving high schools is to be considered a viable strategy for improving student performance, it is important to determine the extent to which differences in schools contribute to differences in student outcomes. This is a fundamental question that Coleman addressed in his landmark study. Colemanôs finding, that no more than a third of achievement differences among students can be attributed to differences in the schools they attend, has been confirmed by virtually every statistical study of school effects since that time. Similarly, a recent international study of student achievement by the OECD found that, on average, differences between schools accounted for 36 percent of the average between-student variation in reading literacy achievement of 15-year-olds among the 26 countries that participated in the study, including 35 percent for the United States (Organization for Economic Co-operation and Development, 2001, p. 60). But the study also found that percentage of variation in student performance attributed to schools varied widely among countries, which suggests that between-school differences in student achievement could be reduced.

These differences may overstate the effects of schools on student achievement. Another of Colemanôs findings was that ñthe social composition of the student body is more highly related to achievement, independent of the studentôs own social background, than is any school factor,ò including school facilities and attributes of teachers (Coleman, 1990, p. 119). This finding too has been replicated in more recent studies. For example, one recent study found that the social class background of the student body had almost as largeð and sometimes largerð an effect on student learning in high school as studentsôown social class background (Rumberger & Palardy, 2005a). The same international study cited earlier found that in the United States the effects of school socioeconomic status (SES) on student achievement (not achievement growth) were about twice as large as the effects of individual SES (Organization for Economic Co-operation and Development, 2001, p. 199). These studies suggest that the widespread segregation of U.S. schools may be a significant contributor to between-school differences in student achievement (Orfield & Lee, 2006).

Another dimension of school performance is the extent to which schools attenuate the effects of individual SES. Although SES is the single most powerful predictor of student achievement among all nations, the extent to which it impacts student achievement varies widely. The same international study found six nationsð Canada, Finland, Iceland, Japan, Korea, and Swedenð that had high levels of student performance and high levels of equality among socioeconomic groups, whereas three nations, including the U.S., had only average levels of student performance and above-average levels of inequality among socioeconomic groups (Organization for Economic Co-operation and Development, 2001, p. 191). These results show that it is possible to have high levels of student performance while still achieving a high level of equality among socioeconomic groups. Within the U.S., studies have found that Catholic schools

similarly have high overall levels of student achievement and more equality among socioeconomic groups (Lee & Bryk, 1989; Bryk, Lee, & Holland, 1993).

One last aspect of school performance concerns the ability of schools to achieve high levels of performance on multiple student outcomes. One recent study found that high schools that were good at raising student achievement were not necessarily good at reducing dropout rates (Rumberger & Palardy, 2005b). Moreover, some attributes that contributed to improved student achievement, such as larger schools, also contributed to higher dropout rates. Accountability provisions to raise student achievement could also have negative effects on other student outcomes, such as dropout rates, if accountability for those other outcomes are weak or non-existent (Rumberger, 2008).

Approaches to Improving High School Performance

Although there are a variety of specific approaches to improving high school performance, they fall into four basic categories.

The first category represents *targeted* approaches, which involve strategies for improving particular facets of the school. The three most common facets are *instructional interventions* designed to improve classroom teaching and student academic achievement in particular subject areas, such as reading or math, or across all areas of teaching; *student support programs* designed to serve students who may need extra support, such as dropout prevention programs to help students who are struggling in school; and *school restructuring* designed to alter the organizational structure of the school, such as restructuring the school into small learning communities where a small group of students and teachers stay together for most or all of the

school day, sometimes over several years. Targeted approaches generally attempt to alter one particular facet of the school, rather than all facets of the school, under the assumption that at least some aspects of the school are functioning adequately, at least for most students, so that reforming a single facet of the school may be sufficient to bring about the desired improvement in school performance.

The second category represents *comprehensive* approaches, which involve altering all aspects of the school, under the assumption that the school itself is not performing adequately for most students, such that targeted approaches would be insufficient to bring about the substantial improvement in student outcomes and high school performance. Two strategies within this category are *reforming existing high schools* by developing a comprehensive set of practices and programs locally or by adopting an externally developed comprehensive school reform model; *creating new* schools, by either developing a new school locally or by adopting an externally developed whole school model. Both of these strategies can be used to create comprehensive high schools to serve all students or to create specialized high schools to serve specific populations of students, such as students who are at risk or have already dropped out of school.

The third category represents *collaborative* approaches, which involve attempts to create partnerships between schools and local government and community agencies, under the assumption that schools alone cannot successfully address the needs of students and improve their performance in school. Instead, schools are more likely to be successful if they work with outside agencies to provide more support to students, and to all the institutionsð families, schools, communitiesð that serve them.

The first three categories all represent school reform strategies, because schools are the focus of change. The fourth category represents *systemic* approaches, which involve making

changes to the entire state educational system, under the assumption that such changes can transform how all schools function in the system, a sort of ñsystemic school reformò (Smith & OõDay, 1991). State and federal accountability systems represent the one recent example of systemic reforms (Furhman & Elmore, 2004).

These approaches have been and continue to be used by schools, districts, state and local governments, and private groups to improve the performance of all schools, including high schools. But in several respects, reforming high schools is more challenging than reforming elementary and middle schools. For one thing, high schools are generally larger than elementary and middle schools.[5] One of the reform principles for high schools, as illustrated below, is to create smaller, more personalized learning environments for students and teachers. A reform strategy is to restructure large schools into small learning communities or create new, small schools, both of which are difficult to implement. Middle and high schools are also difficult to reform because teachers are trained and credentialed in specific subjects, and many schools have departmental structures, both of which make it difficult for teachers to collaborate on instructional reform (Concoran & Silander, 2009). Finally, students vary greatly in their academic preparation and are more challenging to teach, which requires a broad array of instructional offerings and student support services. For example, teachers report that student tardiness, absenteeism, and apathy are much more prevalent among secondary students compared to elementary students (Snyder et al., 2009, Table 72). These challenges have increased over time as all levels of schooling have seen an increase in student diversity, particularly second-language learners.[6]

[5] The average size of high schools in the U.S. was 876 students in 2006-07, compared to an average size of 593 students for middle schools and 446 for elementary schools (Hoffman, 2009, Table 5)

[6] In 2007, more than 20 percent of children ages 5-17 spoke a language other than English at home, more than twice the percentage as in 1979 (Planty et al., 2009, Table A-8-1).

Targeted Approaches to Reform

This section describes a series of specific types of high school reform approaches within each of the four categories. The first three represent targeted approaches.

Instructional improvement programs

The most common approach for improving high school performance is to provide programs for improving classroom instruction. Such programs include: new curricula for teaching academic subjects in reading, math, science, and social studies; new approaches for teaching academic and vocational skills through career and technical education courses; new instructional approaches for teaching the curricula, including the use of technology; and professional development and teacher support programs, such as instructional coaches, mentor teachers, and induction programs for new teachers.

Student support programs

Other programs are designed to provide academic and social support for students to improve their academic performance or help prevent them from dropping out of school. Academic support programs include tutoring and instructional software to supplement classroom instruction. Dropout prevention programs include providing a variety of social services to students and their families to keep students engaged in school. College preparation programs provide additional support to help prepare students for college. These programs typically operate within the comprehensive high school as an add-on or supplement to the studentsôregular school programs.

Another approach to providing support for students is to create separate programs or schools to serve special populations, such as students with disabilities or students at risk for dropping out of school. Instead of serving such students in the regular school setting with supplemental services, alternative education programs are established within regular, comprehensive high schools or in separate schools to serve such students.

The most long-standing example of alternative programs is the continuation high school. Continuation high schools were first established during the expansion of public high schools in the early part of the twentieth century as they began serving an increasingly diverse student body that included recent immigrants, school-age mothers, students with disabilities, students who needed to work, truants, and students with discipline problems. The original continuation high schools were designed to provide a part-time education to students who were working, but gradually they took on a new role in ñadjustmentò education for students considered to be maladjusted for full-time school (Kelly, 1993).

Today, continuation schools are just one type of school within an increasingly widening array of educational options for high school students. The federal government, which collects a variety of data from states on public schools, defines alternative education as:

> A public elementary/secondary school that (1) addresses needs of students that typically cannot be met in a regular school, (2) provides non-traditional education, (3) serves as an adjunct to a regular school, or (4) falls outside the categories of regular, special education, or vocational education (Hoffman, 2009, p. B-1)

In 2006-07, there were 6,638 alternative schools in the U.S. public school system, representing less than six percent of 98,793 public schools and enrolling only about one percent of the students (Hoffman, 2009, p. B-1). But alternative schools enroll about 2.5 percent of all high

school students, with the percentage varying widely among states.[7] In several states (Arkansas, California, Idaho, Minnesota, and Washington) more than five percent of high school students attend alternative schools. A 2002 study found that 48 states had legislation on alternative education, schools, or programs (Lehr, Lanners, & Lange, 2003), although the latest federal government survey found only 40 states that reported high school students enrolled in alternative education schools.

California has an extensive array of alternative schools, including 522 continuation schools, 203 community day schools, and 101 special education schools. In total there were 1,154 alternative high schools in California in 2005-06, compared to 1,037 regular, comprehensive high schools (Rotermund, 2007). Despite the large number of alternative high schools, these schools enrolled only eight percent of all students; yet they accounted for one-third of all the stateôs dropouts.

Restructuring

Another targeted approach to improving high school performance is school restructuring. School restructuring lacks a well-specified, consistent definition, but it generally refers to altering the organizational structure of the school.[8] Some specific examples include:

- Site-based management that provides formal authority to schools through local councils of teachers, administrators, and parents;
- Small learning communities (also referred to as academies or schools-within-schools) composed of a small group of 300-500 students and a group of teachers who stay together over an extended period of time, often for the entire four years of high school.

[7] These figures are based on data from the U.S. Department of Education, Common Core of Data, Build a Table website, retrieved March 14, 2009, from: http://nces.ed.gov/ccd/bat/index.asp

[8] ñOrganizational structures can be defined as roles, rules, and relationships that influence how people work and interact in an organizationò (Newmann, 1993, p. 12).

Some school restructuring activities are schoolwide, while others may simply pertain to parts of the school, such as creating academies for ninth-grade students. Schools may undertake a number of different restructuring activities simultaneously, and sometimes may combine them with other targeted strategies, such as instructional or student support programs.

The next two approaches represent comprehensive approaches.

Comprehensive Approaches to School Reform

Another widespread approach to improving high school performance is known as comprehensive school reform (CSR).

Reforming Existing High Schools

Comprehensive School Reform involves multiple strategies to alter all facets of a school and ñis built on the premise that unified, coherent, and integrated strategies for improvement, knitted together into a comprehensive design, will work better than the same strategies implemented in isolation from each otherò (U.S. Department of Education, 2002, p. 1). The federal governmentôs comprehensive school reform program identifies 11 required program elements (U.S. Department of Education, 2002, p. 1):

(1) Proven methods and strategies based on scientifically based research;

(2) A comprehensive design with aligned components;

(3) Ongoing, high-quality professional development for teachers and staff;

(4) Measurable goals and benchmarks for student achievement;

(5) Support within the school by teachers, administrators, and staff;

(6) Support for teachers, administrators, and staff;

(7) Meaningful parent and community involvement in planning, implementing, and evaluating school improvement activities;

(8) High-quality external technical support and assistance from an external partner with experience and expertise in schoolwide reform and improvement;

(9) Evaluation of strategies for the implementation of school reforms and for student results achieved, annually;

(10) Resources to support and sustain the schoolôs comprehensive reform effort;

(11) Strategies that have been found to significantly improve the academic achievement of students or that have strong evidence to suggest they will significantly improve the academic achievement of students.

Although individual schools may undertake comprehensive school reform, the widespread interest in CSR has come from the development of specific design models that can be replicated or scaled up in other sites. The idea behind this approach is that instead of countless individual schools attempting to identify and implement research-based strategies on their own, external developers would design and package a comprehensive set of proven strategies in a coherent model and then assist other schools in implementing their model with financial support from the government or the private sector. A review from 2003 identified 33 CSR models that had been replicated in 10 or more schools (Borman et al., 2003).

New Schools

The other comprehensive strategy is to create new schools. This strategy is based on the idea that it may be easier to create a new school than to reform an existing one. The strategy is often coupled with creating a particular type of new school known as a charter schoolð public schools that are established and managed outside the regular public education system, and that are freed from most of the regulations and requirements of regular public schools. Charter schools were established to provide choice within the public school system, to spur educational innovation, to

create competition as a way to improve non-charter schools, and to improve student achievement. The extent to which charter schools have achieved those goals has been the subject of intense and often partisan debate (Carnoy, Jacobson, Mishel, & Rothstein, 2005; Finn, Manno, & Vanourek, 2000; Henig, 2008; Zimmer et al., 2009). Nonetheless, their popularity continues to grow. In 2006-07 there were 4,132 charter schools in the U.S., up from 2,179 in 2003-04, with 521 of those secondary charter schools (Snyder et al., 2009, Tables 98, 101). In California alone, the number of charter high schools increased from 97 in 2000-01 to 271 in 2005-06 (Rotermund, 2007).

Collaborative Approaches

Collaborative approaches involve establishing collaborative relationships between schools and outside organizations, such as government agencies and local community organizations. This strategy is based on the idea that schools do not have the resources or expertise to attend to all the needs of their students and their studentsô families.

> The primary institutions that serve youthð health, schools, employment, trainingð are crucial and we must begin with helping them respond more effectively to contemporary adolescent needs. Effective responses will involve pushing the boundaries of these systems, encouraging collaborations between them and reducing the number of adolescents whose specialized problems cannot be met through primary institutions (National Research Council, Panel on High-Risk Youth, 1993, p. 193).

Outside groups can provide a range of services, such as medical and counseling services, and after-school and summer-school programs. In some cases the services can be provided on the school campus, such as medical and dental clinics, or after-school programs. Partnerships with the business community can provide computers and other material resources, mentors and tutors, and work-related opportunities.

Systemic Approaches

Systemic approaches involve making changes to the entire educational system. The constitutional authority for providing public education lies with states, so each state in the U.S. has developed its own public education system. The state systems provide funds for schools, establish the curriculum, set certification requirements for school personnel, determine performance requirements for students and school, among other things, as well as determine the level of authority given to local educational systems on the operation of schools.

These systemic features generally apply to all levels of schooling in the public system. However, some apply specifically to high schools. One is establishing high school graduation requirements, which we reviewed earlier. Another is setting the compulsory schooling ageð the age to which students must attend school. A third is setting enrollment requirements for colleges and universities that allow high school students to take college courses while still in high school, which is known as dual enrollment (Krueger, 2006). Some dual enrollment programs are established in new types of high schools, known are middle college high schools or early college high schools located on college campuses.

The Difficulty in Evaluating Effective Approaches

To improve high school performance requires determining the effectiveness of specific approaches as well as their critical features. But identifying whether reform strategies are truly effective is actually quite difficult. Although there is an extensive body of research that has been conducted on high schools, little of it can identify whether the school itself or the programs within it are effective.

One popular method is to conduct case studies of high schools that have somehow been identified as effective or high performing. For example, one recent case study of five high schools based selection criteria on a mixture of school practices, multicultural pedagogy, and a broad array of student outcomes not limited to student test scores (Friedlaender & Darling-Hammon, 2007). Case studies can provide rich and detailed descriptions of the origins, practices, and outcomes of schools, but they cannot by themselves determine whether the school is actually effective in producing good outcomes for students. The reason is that other factors could account for the schoolsô apparent success. For example, many apparently successful schools, particularly charters and magnet schools, require students and parents to choose the school and perhaps fill out an application. They may also require some amount of parental involvement, such as volunteering in the school. In some cases, the school also selects the students based on criteria such as past performance or commitment to the school requirements. These so-called ñselection effectsò can result in a student body that is different from the student bodies of other schools. Thus, case studies are unable to determine whether the outcomes of the school are due to the characteristics of the school or the characteristics of the students.

Another method for studying high schools is to use statistical models to test the relationship between student outcomes, such as test scores and dropout rates, and a variety of student and school characteristics. These studies are often based on national longitudinal studies conducted by the federal government involving large, national samples of students and schools, and a broad array of data based on student, parent, teacher, and administrator surveys, student test scores, and institutional data, such as student transcripts. And because the data are collected over a number of years, the studies can be used to examine changes in student outcomes over time. One of the most widely used studies is the National Educational Longitudinal Study of

1988 (NELS: 88), a study of 25,000 eighth-graders in the spring of 1988 who were surveyed throughout their high school and college careers until 2000, when most participants were 26 years of age. A large number of studies of high school effectiveness have been conducted using the NELS:88 data (Bryk et al., 1993; Lee & Smith, 1995; Lee & Smith, 1997; Rumberger and Palardy, 2005b).

The statistical models used in these studies control for differences in observed student characteristics to help determine the extent to which differences in student outcomes can be attributed to differences in student characteristics or school characteristics. And because of the broad array of data in the datasets, the studies are able to determine which features of schools predict student outcomes after controlling for other factors. This can help determine the relative importance of various factors. Yet despite the wide array of data collected in these studies, they may not be able to identify all of the factors related to student success, particularly factors that are not measured in the data set. Nonetheless, these studies do provide valuable evidence on the effectiveness of high schools.

The most rigorous evidence on the effectiveness of high schools comes from evaluation studies. Evaluation studies are used to study the effectiveness of a wide variety of interventions, from single interventions, such as small classes, to comprehensive school reform (CSR) models. There are a variety of research designs for conducting evaluation studies, and the rigor of the design dictates the ability of determining a causal connection between the intervention and student outcomes. The so-called ñgold standardò in evaluation studies is the randomized experiment, more formally referred to as the *randomized controlled trial* (RCT), where students are randomly assigned to either the intervention (experimental group) or the regular or non-reform program (control group). For single interventions or programs, it is possible to randomly

assign students to the treatment or control condition. One well-known, large-scale example is the Tennessee class-size reduction study, where students in kindergarten to third grade were randomly assigned to small (15 students per class) or regular (25 students per class) classes (Finn & Achilles, 1999).

Because it is virtually impossible to randomly assign students to schools, an alternative design for evaluations of CSR models is to randomly assign reform models to schools (Cook, 2002). But this technique is costly and difficult, requiring a large sample of schools in order to establish strong causal inferences (Raudenbush, Martinez, & Spybrook, 2007). A more common evaluation design is the quasi-experimental design that is not based on random assignment, but instead uses statistical techniques to control for differences in the characteristics of students attending experimental and control schools. Several techniques can be used to estimate causal effects from quasi-experimental studies (Schneider et al., 2007).

Evaluating the effectiveness of new schools, such as charters, presents special challenges (Zimmer et al., 2009, pp. 21-26). Randomized evaluations can only be conducted in schools that are over-subscribed, allowing comparisons between admitted and non-admitted students, but such evaluations can only determine whether those particular schools are effective, not if other schools of the same general type (e.g., all charter schools) are. Alternatively, evaluations that are more rigorous can be conducted by comparing the achievement gain trajectories of students who transfer into new schools to students in traditional public schools, using longitudinal achievement data, although this approach requires a number of assumptions to establish a causal connection (pp. 24-25).

What is Known About the Effectiveness of High School Reform Approaches

There is an increasing awareness of the need to base educational reform efforts on rigorous, scientific evidence (Slavin, 2008). The No Child Left Behind (NCLB) Act, for instance, recommends programs and practices ñbased on scientifically-based researchò more than 100 times (p. 5). But evidence that counts as ñscientifically basedò is open to question and has led to inconsistent conclusions from various efforts by the federal government and others to review and synthesize the existing research evidence on education reform strategies.[9] Differences in the conclusions result largely from the standards used to determine whether a particular study should be included in the review. Nonetheless, virtually all reviews find a paucity in both the number and quality of existing studies.

In 2002, the U.S. Department of Education established the What Works Clearinghouse (WWC) to review scientific evidence on the effectiveness of a variety of educational interventions.[10] As of June 2009, WWC has published only one evaluation of an adolescent literacy program, and of the 36 studies of that program, only three meet the WWC evidence standards for scientific rigor. The WWC reviewed 84 studies of 22 dropout prevention programs through September 2008, and found only 23 studies of 16 interventions that met their evidence standards. An earlier U.S. Department of Education center, the Comprehensive School Reform Quality Center, reviewed 197 studies of 18 secondary comprehensive school reform models and found only 28 studies of 16 models that met their standards (Comprehensive School Reform Quality Center, 2006). An independent scholarly review of 232 studies of 29 CSR models found only 109 studies conducted by third-party evaluators using comparison group designs, and only 7

[9] See the January/February 2008 issue of *Educational Researcher.*
[10] See the What Works Clearinghouse website at: http://ies.ed.gov/ncee/wwc/

studies of three CSR models that were based on randomized experiments, the so-called gold standard (Borman et al., 2003, p. 163).

Results from rigorous evaluations can determine not only whether a particular intervention is effective, but the magnitude of the effect, known as the effect size (ES). Although there are no absolute standards for judging the magnitude of effect sizes, one prominent statistician argues that an ES of at least .2 should be considered a small effect (which corresponds to increasing the likelihood of graduating from 50 to 58 percent); an ES of at least .5 should be considered a medium effect (which corresponds to increasing the likelihood of graduating from 50 to 69 percent); and an ES of at least .8 should be considered a large effect (which corresponds to increasing the likelihood of graduating from 50 to 79 percent) (Cohen, 1988, pp. 25-26). To illustrate, an evaluation of the Tennessee class size experiment found that students who were enrolled in small classes from kindergarten through third grade had high school graduation rates that were 11 percentage points higher than students who were enrolled in regular-sized classes, which is an effect size of about .25 (Finn, Gerber, & Boyd-Zaharias, 2005). The effects were even stronger for low-income students; those in small classes had graduation rates 18 percentage points higher, which is an effect size of .50. A review of 232 evaluation studies of the effects of 29 different CSR models on student test scores found an average effect size of .12, although interventions that had been implemented for eight years or longer had an average effect size of .50 (Borman et al., 2003).

Although evaluation studies are able to establish a causal connection between the overall intervention and student outcomes, the evaluations are generally not able to determine the specific causal mechanisms responsible for the outcome. Comprehensive strategies, such as those found in CSR models, typically involve a series of reform components, from structural

features (such as creating small learning communities), to specific instructional components. Consequently, it is impossible from a whole-school evaluation to determine which components are critical to the modelsô effectiveness. Some components may be critical and others not, but the evaluation is unable to make this determination unless it is implemented in such a way that the effectiveness of specific components can be determined.

Evaluations of specific reform strategies are generally confined to externally developed programs or school reform models. One reason is that external developers have an incentive to determine the effectiveness of their strategies so that they can establish their legitimacy and convince other schools and districts to adopt their strategies. Thus, developers will conduct their own evaluations or help raise funds to have third-party, independent organizations evaluate their programs or models. Local schools and districts have less incentive or expertise to conduct scientifically rigorous evaluations of their strategies, although they should do so in order to determine whether the strategies are indeed effective.

The existing reviews of all school reform strategies, regardless of school level, find that relatively few reform strategies have been evaluated with rigorous, scientific methods, even when acknowledging the different standards used to determine rigor. There are even fewer evaluations of reform strategies specifically focused on high schools. But there are some reviews of strategies in both middle and high schools, which are discussed below.

Instructional Improvement Programs

Three recent reviews were conducted on instructional programs to improve the achievement of middle and high school students. One review examined 100 studies of three types of programs designed to improve achievement in mathematics (Slavin, Lake, & Groff, 2009). In this review,

40 studies of mathematics curricula found very small effects (ES = +0.03); 38 studies of computer-assisted instruction found small effects (ES = +0.10); and 22 studies of instructional process programs found small effects (ES = +0.18); although the effects of specific programs varied widely, with studies of two forms of cooperative learning having medium effects (ES = +0.48). An earlier review examined 33 studies of four types of programs designed to improve achievement in reading (Slavin, Cheung, Groff, & Lake, 2008); Regarding these programs, no studies of secondary reading curricula met the criteria to be included in the review; eight studies of computer-assisted instruction found small effects (ES = +0.10); 16 studies of instructional-process programs had small effects (ES = +0.21); and nine studies of two mixed-method models that combined large-group, small-group, and computer-assisted, individualized instruction had small effects (ES = +0.23). The third review was conducted by the What Works Clearinghouse based on three studies of a computer-based adolescent literacy program that supplements regular classroom reading instruction in grades K-8.[11] The review found that the program had small effects on reading comprehension (ES = .27) and literacy achievement (ES = .28).

Overall, instructional programs that focus on daily teaching methods and student interactions, such as cooperative or group learning methods, appear to have more impact on student achievement than curricula or computer-based approaches (Concoran & Silander, 2009).

In addition to conducting reviews of scientifically rigorous research studies on specific instructional programs, the What Works Clearinghouse also issues Practice Guides that use expert panels to consider a wider range of research evidence and to recommend a set of practices or strategies for addressing particular educational problems and some suggestions on how to carry them out.[12] One recent report made five recommendations on effective practices to

[11] See: http://ies.ed.gov/ncee/wwc/reports/topic.aspx?tid=15
[12] See: http://ies.ed.gov/ncee/wwc/publications/practiceguides/

improve adolescent literacy in middle and high schools that the panel considered to have moderate to strong levels of supporting evidence (Kamil et al., 2008, p. 7):

1. Provide explicit vocabulary instruction.
2. Provide direct and explicit comprehension strategy instruction.
3. Provide opportunities for extended discussion of text meaning and interpretation.
4. Increase student motivation and engagement in literacy learning.
5. Make available intensive and individualized interventions for struggling readers that can be provided by trained specialists.

Student Support Programs

The most common student support programs in high schools are generally those designed to keep at-risk students from dropping out. The What Works Clearinghouse reviewed 23 qualifying studies of 16 dropout prevention interventions in middle and high schoolsð 12 of the programs were student-support programs and four were CSR or new school modelsð and assessed their effectiveness in improving three student outcomes: (1) staying in school; (2) progressing in school, and (3) completing school (U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2008). Of the 12 student support programs, 5 were effective in keeping students in school, 4 were effective in helping students progress in school, and 4 were effective in helping students to complete school, although none of the 4 programs were effective in helping students earn a regular high school diploma.

The What Works Clearinghouse also issued a Practice Guide on dropout prevention that provided six recommendations for improving dropout and graduation rates that were supported by low or moderate levels of evidence (Dynarski et al., 2008, p. 6):

1. Utilize data systems that support a realistic diagnosis of the number of students who drop out and that help identify individual students at high risk of dropping out.
2. Assign adult advocates to students at risk of dropping out.
3. Provide academic support and enrichment to improve academic performance.
4. Implement programs to improve studentsô classroom behavior and social skills.
5. Personalize the learning environment and instructional process.
6. Provide rigorous and relevant instruction to better engage students in learning and provide the skills needed to graduate and to serve them after they leave school.

Another type of student support program helps prepare underrepresented students for college by offering additional services, such as mentoring, academic support, and a peer support network. A review of such programs found that few had been subjected to rigorous evaluations and that, in general, the programs had little impact on academic achievement (Gándara & Bial, 2001).

Restructuring Programs

A recent review of research on smaller learning communities found mixed results: ñsome studies suggest encouraging findings about the benefits of SLCs on student achievement outcomes, while others suggest mixed or even negative results (Fleischman & Heppen, 2009, 119). In general, reform efforts are more successful in improving student engagement-related outcomes, such as attendance and promotion rates, rather than academic achievement. But as the authors point out, it is difficult to attribute outcomes to the creation of smaller learning communities alone, since restructuring is almost always accompanied by other reform strategies. An evaluation of the federal Smaller Learning Communities Program is reported below.

Comprehensive School Reform

Several reviews have been conducted on the effectiveness of CSR models. The Comprehensive School Reform Quality (CSRQ) Center reviewed 28 studies of 18 comprehensive school reform models: 10 of the models were found to produce positive effects in student achievement, but only 3 were found to provide positive effects in dropout and graduation rates (Comprehensive School Reform Quality Center, 2006). An independent review of 10 comprehensive school reform models where at least half of the reform focused on secondary schools found only two evaluations conducted by outside evaluators that showed significant improvements in student outcomes (Borman et al., 2003). And the What Works Clearinghouse review of four CSR or new school models found that only one was effective at keeping students in school, two were effective in helping students progress in school, and none were effective in helping students to complete school (U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2008).

The What Works Clearinghouse also issued a Practice Guide on turning around low-performing schools that made four recommendations based on only low levels (as defined by expert opinion) of evidence (Herman et al., 2008, p. 8):

1. Signal the need for dramatic change with strong leadership.
2. Maintain a consistent focus on improving instruction.
3. Make visible improvements early in the school turnaround process (quick wins).
4. Build a committed staff.

New Schools

If one strategy for improving high school performance is to close, rather than redesign, low-performing schools and replace them with new schools, then it is important to consider the evidence about the effectiveness of new schools. The biggest effort in this regard is the creation of charter schools. Although there have been an increasing number of studies of charter schools, many reviews of these studies have been conducted by persons or institutions with strong, public views about their efficacy (Henig, 2008), making it difficult to identify unbiased reviews of the research evidence.

Two large studies of charter schools have recently been completed. The first study by RAND, an independent research organization, examined 669 charter schools in five districts located within five different states and statewide in three other states (Zimmer et al., 2009). The study found (pp. xii-xv):

1. No systematic evidence to support the fear that charter schools are ñskimming offò the highest-achieving students.

2. In five locales, middle and high school charters are producing achievement gains that are, on average, neither substantially better nor substantially worse than those of traditional public schools (TPS), while in Chicago (reading) and in Texas (in both reading and math) charter middle schools are falling short of traditional middle schools.

3. In most locations, charter schools have difficulty raising student achievement in their first year of operation, typically producing achievement results that fall short of local TPS.

4. Charter schools in most locales have marginally greater variation in performance than TPS.

5. In the two locations with attainment data (Florida and Chicago), attending a charter high school is associated with statistically significant and substantial increases in the probability of graduating and attending college.

6. There is no evidence in any of the locations that charter schools are negatively affecting the achievement of students in nearby TPS; but there is also little evidence of a positive competitive impact on nearby TPS.

The study goes on to point out that one of the most important implications is the need ñto examine a broad and deep range of student outcome measures and to provide evidence on the mechanisms producing positive long-term impactsò (p. xix).

The second study by the Center for Research on Education Outcomes (CREDO) at Stanford University examined Charter school performance in 15 states and the District of Columbia (Center for Research on Education Outcomes, 2009). The study found (pp. 3-6):

1. Forty-six percent of the charter schools had math gains that were indistinguishable from the average gains among matched students who attended traditional public schools (TPS), 17 percent had gains that exceed the growth in TPS, and 37 had gains significantly below those in TSP;

2. The effectiveness of charter schools varied among states and was affected by the charter school policies under which the schools operated;

3. Charter school students in elementary and middle school grades had significantly higher rates of learning than their peers in TPS, while students in charter high schools and multilevel schools had significantly worse rates of learning;

4. Charter school results varied by family backgroundð students living in poverty and English language learners do significantly better; Black and Hispanic students do significantly worse;

5. Students in charter schools do better over timeð first year students experience a decline in learning, while students in years two and three show positive gains.

Although most charters are locally developed, there are some well-known and widely acclaimed externally developed charter school models. Two of the best known are the Knowledge Is Power Program (KIPP), which began in 1994 and now comprises a national network of almost 66 public schools (85 percent middle schools) in 19 states and the District of Columbia;[13] and the Green Dot Public Schools, which began in 1999 with five charter schools in the Lennox and Los Angeles Unified School Districts and is in the process of transforming two large high schools into a number of smaller charter high schools, and will soon open one in New

[13] See: http://www.kipp.org/

York.[14] A recent, rigorous evaluation of five KIPP middle schools in the San Francisco Bay area found that KIPP students outperformed a matched comparison group during the school year, although the evaluation only controlled for observed, not unobserved, differences between KIPP and comparison students; moreover, some of the KIPP schools experienced high student attrition, which could also affect student performance comparisons (Woodworth, David, Guha, Wang, & Lopez-Torkos, 2008).

One unique high school reform effortð known as Early College Schools (ECS)ð represents a combination of strategies. The schools are created from partnerships with two- or four-year colleges through conversion of existing high schools, or as new schools, or as new programs within existing schools. Students take both high school and college courses and earn college credit along with their high school diploma. In the fall of 2006, there were 130 ECS located in 21 states and the District of Columbia. The schools are supported by the Early College High School Initiative with funding provided by the Bill & Melinda Gates Foundation (Berger et al., 2008). To date, no outcome evaluations have been completed.

Summary of Evidence on School Reform Strategies

In summary, there is relatively little scientifically rigorous evidence about the effectiveness of specific reform strategies for improving high school performance. Moreover, there are discrepancies among reviewers of the scientific evidence, which may create confusion among educators who wish to understand and use scientific evidence in pursuing high school reform strategies. Yet there are examples of effective programs and models within the various typesð instructional strategies, dropout prevention programs, comprehensive school reform models, and charter schools. No type of strategy is inherently more effective than the others, although it

[14] See: http://www.greendot.org/

seems logical that more comprehensive strategies should have the potential to improve high school performance to a greater degree than targeted strategies, simply because they involve improving more facets of the school. There is also some limited evidence to support this claim (Glennan Jr., Bodilly, Galegher, & Kerr, 2004, p. 17). This may be especially true if high school performance is measured with multiple indicators, such as test scores as well as graduation and dropout rates. So, while instructional programs may improve student achievement, and dropout prevention programs may improve high school graduation rates, comprehensive strategies can improve both sets of outcomes. Finally, the existing evidence suggests that neither comprehensive strategy for improving high school performance (redesigning existing schools versus starting new schools) is inherently superior to the other.

Systemic Reform Strategies

Three systemic reforms have attempted to improve high school performance. One is to raise high school graduation requirements. As discussed earlier, an increasing number of states have raised the requirements for high school graduation, both by increasing the number and rigor of courses needed to receive a diploma and by requiring students to pass a high school exit exam. A number of studies have been conducted evaluating the impact of high school exit exams. A recent review of these studies found that while earlier studies found limited or no effects of such exams, more recent studies have found that high school exit exams have lowered high school completion rates, especially among low-achieving students (Reardon, Atteberry, Arshan, & Kurlaender, 2009; Rumberger & Lim, 2008; Warren et al., 2006).

Another systemic reform is to raise the compulsory schooling age. In some states the maximum schooling ageð the age at which students no longer have to attend schoolð is 16 or

17, which means students do not have to stay in school long enough to graduate (Snyder et al., 2009, Table 165). One policy recommendation for improving graduation rates is to raise the maximum compulsory schooling age to 18 (Bridgeland, DiIulio Jr., & Streeter, 2008). A recent review of several studies that examined the relationship between the state compulsory schooling age and dropout or graduation rates found that states with higher compulsory schooling ages had lower dropout rates and higher graduation rates (Rumberger & Lim, 2008).

A third systemic reform strategy to improve high school performance is establishing dual-enrollment programs, which allow high school students to take college classes while still completing their high school diploma. In 2003, 71 percent of U.S. high schools offered dual-enrollment courses, with more than 1 million students enrolled (Snyder et al., 2009, Table 152). No rigorous evaluations have been done of dual-enrollment programs within high schools, but a rigorous evaluation was conducted of a middle college program in Seattle that was part of the federally funded dropout prevention program reviewed below. The program was not found to improve dropout or graduation rates (Dynarski & Gleason, 1998). As mentioned earlier, the Gates Foundation is supporting a dual-enrollment program known at the Early College High School Initiative, but no summative evaluations have been completed (Berger et al., 2008).

Past Initiatives to Improve High School Performance

Numerous large-scale initiatives have been undertaken in the past 20 years to improve the performance of high schools in the U.S. These efforts have been undertaken by government agencies, foundations and non-profit organizations, state and local governments, and independent developers. Some of these efforts have focused on all levels of schooling, while

others have focused specifically on high schools. Some efforts have attempted to improve existing schools by supporting locally developed reforms or externally developed programs, while others have promoted the development of new schools. Yet despite these extensive and widespread efforts and the huge sums of money spent, they have yielded relatively few significant and lasting improvements in high school performance. This section describes the largest and most extensive efforts and their impact.

Federal Programs

The federal government has supported high school reform since the initial passage of the Smith-Hughes Act in 1917, which provided federal aid to vocational education in high schools. Most of the support for high school reform has been subsumed in reform efforts targeting all levels of K-12 education, but some efforts have specifically targeted high schools. **Table 2** provides a summary of the largest and most recent efforts.

Vocational Education

The purpose of vocational education is ñto develop more fully the academic and career and technical skills of secondary and postsecondary students who elect to enroll in career and technical programs.ò[15] The program provides funding to states for state leadership activities, administration of the state plan for vocational and technical education, and sub-grants to eligible recipients to improve vocational and technical education programs, including the use of technology and professional development. The program is funded through the Carl D. Perkins Vocationalï Technical Education Act Amendments of 1998 (Perkins Act), Title I (refunded in

[15] See program description at: http://www.ed.gov/programs/ctesbg/index.html

2006) and administered through the federal Office of Vocational and Adult Education. The program provides about $1.1 billion in funding to states per year.

A comprehensive review of the status of vocational education and the impact of the 1998 Perkins Act was completed in 2004 (U.S. Department of Education, Office of the Under Secretary, Policy and Program Studies Service, 2004). The review examined the value-added effects of high school vocational education on a number of student outcomes: academic achievement, high school completion, postsecondary enrollment and completion; and short- and medium-run earnings. The most consistent positive finding was that students who completed an extra vocational course earned about three percent more in their jobs after completing school than students without the extra course. But there was little evidence that vocational courses improved *academic* outcomes:

> Both analyses of high school student data and randomized controlled studies indicate that, on average, vocational courses and programs do not themselves ñadd valueò to academic achievement as measured by test scores. Not surprisingly, substituting additional academic courses for occupational courses does raise achievement. Moreover, although there is mixed evidence that vocational education reduces dropping out of school, the more rigorous studies suggest there is no effect (p. 7)

The review concluded that the quality improvement strategies in the 1998 revisionsð which included the integration of academic and vocational instruction, linkages between secondary and postsecondary programs, collaboration with employers, and the use of technologyð were perhaps too vague to drive meaningful change; moreover, little is known about the effectiveness of these strategies (p. 9).

The reauthorization of the Perkins Act in 2006 included four significant changes to improve its impact:

> (1) States must create a series of new career and technical education (CTE) offerings called ñCTE Programs of Study,ò coherent programs of study aligned with postsecondary programs and credentials;

(2) Local recipients are held accountable for student achievement in a series of performance indicators, including high school graduation;

(3) States can maintain existing Tech Prep programsð early, rigorous forms of high school CTE linked to postsecondary studies and apprenticeshipsð with increased accountability, or they can merge them into the general grant program;

(4) The new purpose of the Perkins Act is to ñprovide individuals with opportunities throughout their lifetimes to develop, in conjunction with other education and training programs, the knowledge and skills needed to keep the United States competitiveò and to encourage preparation for ñhigh-skilled, high-wage and high-demandò careers (Meeder, 2008).

Whether these revisions address the shortcomings found in the earlier evaluation remains to be seen.

Comprehensive School Reform Program

Another major effort for reforming public schools comes from the Comprehensive School Reform (CSR) Program, which was first authorized as a demonstration program in 1998 and then authorized as a regular program in 2002 as part of the No Child Left Behind Act (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service., 2008a, p. xiii). The federal CSR program shares the same premise of all CSR reform strategies outlined earlier: it attempts to provide coherent improvements to all aspects of a schoolôs operation.

The U.S. Department of Education funded two evaluations of the CSR program. The first focused on the CSR demonstration program, where the federal government provided funds for schools to implement an approved list of 17 specific CSR models or to develop their own models. The second evaluation focused on the reauthorized CSR program that provided funds to implement any CSR program that included the 11 essential components (described earlier) and was based on proven research and evaluation methods. Both programs focused on schools

identified as ñhigh-povertyò under Title I of the federal Elementary and Secondary Education law (at least 50 percent of the students were eligible for the federal school lunch program).

The first evaluation, National Longitudinal Evaluation of Comprehensive School Reform (NLECSR), was funded in 2000 to study the adoption, implementation, and impact of eight externally developed CSR models (Accelerated Schools Project, ATLAS Communities, Co-nect, Expeditionary Learning/Outward Bound (ELOB), Modern Red Schoolhouse, Success for All (SFA)/Roots and Wings, Turning Points, and Urban Learning Centers) in 650 elementary and middle schools and 21 districts across 17 states (Aladjem et al., 2006). The NLECSR evaluation found:

- The quality of the adoption processð as measured by how informative, inclusive, and legitimate it wasð varied widely, yet was not related to the fidelity of the modelsô implementation because, as other research has shown, commitment can follow coerced or mandated involvement (McLaughlin, 1990).
- Implementation of model-like practices was not higher in CSR schools compared to matched comparison schools, although there were vast differences across the various CSR models.
- Only 10-20 percent of the schools adopted CSR models comprehensively, that is, adopting all of the components of the models (e.g., governance, instruction, etc.).
- Schools that achieved uniformly high levels of implementation across components and over 3-5 years of time demonstrated significant improvement in reading and math outcomes.
- Although some models were more successful than others in building the collective commitment and professional community of teachers (sometimes referred to as social capital), these features were not related to improvements in student achievement.
- Nearly a third of the schools ended their relationships with the model providers after three years, citing a wide variety of reasons, including low school capacity, insufficient funding, instability of school leadership, faculty turnover, and competing reforms (pp. 2-7).

The evaluation also studied the role that states and districts play in the adoption and support of CSR activities. Building on the existing research literature, the evaluation examined:

(1) the *alignment* between CSR strategies and existing policies, such as curriculum and professional development, (2) the strategies to *support* the adoption and implementation of CSR models, and (3) the *refinement* of CSR policies based on feedback and formal evaluations (p. 46).

The study found that states differed in the extent of alignment between CSR strategies and existing policies. Some states had developed comprehensive support systems that included tools (e.g., data analysis, aligned funding, monitoring implementation) and structures (e.g., school support teams, school improvement coaches), whereas others had no existing system and CSR was simply a funding stream for school improvement (pp. 49-50). States also differed in how they responded to a common problem of school capacity:

> One of the most common and fundamental challenges states encountered in developing their approaches to CSR was the fact that they were dealing with high-poverty, low-performing schools which often lacked the capacity to develop a comprehensive school reform plan that was (1) compatible with their particular needs and (2) capable of being successfully implemented (p. 51).

States provided differing levels and types of technical assistance to help schools develop capacity, some extensive and some minimal. Finally, states differed in the extent which they adapted their CSR policy over time and in response to feedback.

Districts, too, varied in the extent and types of CSR alignment, support, and refinement. Some districts adopted districtwide curriculum or instructional programs and aligned the CSR models to such programs. Some districts integrated CSR activities into other reform strategies, which in some cases meant they controlled which schools implemented CSR and restricted the selection of CSR models to avoid conflicts with other reform efforts. The study also found that districts that tended to be controlling and restrictive in selecting school reform models also tended to provide more support to the schools. Finally, the study found that some districts were

quick to phase out CSR if it was not perceived as effective in raising student test scores, in part because of the pressure from accountability systems like NCLB to quickly produce results.

The second evaluation was a five-year study of the reauthorized CSR program, funded by the U.S. Department of Education in 2002 (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service, 2008a). This ongoing study, conducted by WestEd and COSMOS Corp., focuses on schools nationwide that received their initial CSR awards in 2002. The study compares a random sample of 500 CSR schools and 500 matched non-CSR schools. Because only a small number of high schools received CSR awards in 2002, the study only examines the implementation and impacts of the program on elementary and middle schools. To date, the main findings of this study have been:

- Receipt of a CSR award was not associated with school-level achievement gains in mathematics or reading achievement through the first three years of the award.
- Both CSR and non-CSR schools implemented an average of fewer than 4 (of the 11 required) components in 2003, and fewer than 5 in 2005, at both the elementary and middle school levels.
- Only one-third of 2002 CSR awardees chose reform approaches with recognized scientific research bases.
- The comprehensiveness of CSR reform implementation, as measured by the number of CSR components implemented, was not related to achievement gains in mathematics and reading at the elementary or middle school levels.
- Mathematics achievement improved in low-performing elementary schools, as well as in all middle schools (including low-performing schools), that adopted models with scientific research bases (pp. xvii-xx).

One limitation of the study, as the authors point out, is that the impacts of the reforms on student achievement were measured in the third year of implementation (p. 56), while a recent review of CSR models found the greatest impacts after five years (Borman et al., 2003).

In addition to providing direct funding to states, districts, and schools to implement CSR models, the federal government has also funded two related support programs described below:

Comprehensive School Reform Quality Initiatives

This program provided $21.5 million over the four-year period from 2004-2008 to support eight non-profit organizations, universities, and consortia to:[16]

- Foster the development of CSR models.
- Strengthen effective capacity building for CSR providers to expand and ensure quality of their work.
- Provide direct technical assistance and support to state education agencies (SEAs), local education agencies (LEAs), and schools.
- Develop and disseminate tools and resources to assist planning and implementation.
- Conduct evaluations to assess progress.
- Conduct evidence reviews on the effectiveness and quality of a large number of CSR programs.

Comprehensive Centers Program

This program, initiated in 2005, provides $57 million annually to support two types of comprehensive centers to help increase state capacity to assist districts and schools to meet their student achievement goals:[17]

- Sixteen regional centers that (1) provide services primarily to state education agencies (SEAs) to enable them to assist school districts, regional education agencies, and schools, especially low-performing schools; and (2) disseminate information and reports on improving academic achievement, closing achievement gaps, and sustaining school improvement to schools, teachers, parents, and policymakers.
- Five content centers that focus on five specific areasð accountability, instruction, teacher quality, innovation and improvement, and high schoolsð and supply much of the research-based information and products in the specific area that regional centers will use when working with states.

[16] This information retrieved July 17, 2009, from: http://www.ed.gov/programs/qualinits/index.html
[17] This information retrieved July 17, 2009, from: http://www.ed.gov/programs/newccp/index.html

School Improvement Fund

This program has replaced the Comprehensive School Reform Program. It will be funded at $3 billion under the American Recovery and Reinvestment Act of 2009 (Recovery Act) that was signed into law by President Obama on February 17, 2009.[18] Forty percent of those funds will be targeted to middle and high schools.

While the preceding school improvement efforts focused on all levels of schooling, three additional programs focused exclusively on middle and high schools.

School Dropout Prevention Program

The School Dropout Prevention Program was funded from 2004 to 2006 to support dropout prevention and reentry programs in middle and high schools. The program was preceded by two phases of an earlier program, the School Dropout Demonstration Assistance Program (SDDAP), one operating from 1989 to 1991, the other from 1991 to 1996. Mathematica Policy Research conducted an evaluation of 20 programs around the country that were funded in the second phase of the SDDAP program: eight middle school programs, eight high school programs, and five schoolwide restructuring programs (Dynarski & Gleason, 1998; Dynarski, 2004). All of the programs shared two features: (1) they provided support to students to overcome personal, family, and social barriers to school success; and (2) they tried to create smaller and more personal settings in which students could feel secure and learn more effectively (Dynarski & Gleason, 1998, p. 4).

Of the eight high school programs, five were alternative school programs (four were alternative high schools and one was a school within a school) that offered a regular high school diploma. The evaluation found that none of the five alternative school programs lowered the

[18] Information retrieved July 17, 2009, from: http://www.ed.gov/policy/gen/leg/recovery/index.html

dropout rate. In four of the five sites, more students earned diplomas compared to control students, but the differences were not statistically significant. The other three programs prepared students to take the GED, producing high GED and overall (GED plus diploma) completion rates.

The largest and most ambitious programs were the five school restructuring efforts funded by multimillion dollar grants to reduce the need for alternative schools or programs (Dynarski & Gleason, 1998). The evaluation found that none of these restructured schools significantly reduced dropout rates in relation to comparable schools.

> The evaluation did not observe much change, however, or even signs of it beginning. Restructuring schools found it easier to add dropout-prevention services than to change teaching and learning. Some initiatives managed to change teaching and learning to a degree, but the changes were fragile and easily undone if district leadership changed or local political contexts shifted (p. 14).

The study found that there was little consensus about the source of the dropout problem and, in particular, how faculty and staff may have contributed to it. Consequently, few faculty and staff members were eager or willing to change what they were doing. Finally, turnover of district administrators undermined support for change.

Smaller Learning Communities Program

Another high-school-focused program was the Smaller Learning Communities Program, which provided more than $1 billion over a nine-year period to school districts to support the implementation of smaller learning communities (SLCs). The SLCs included structures such as freshman academies, multi-grade academies organized around career interests or other themes; ñhousesò in which small groups of students remain together throughout high school; and autonomous schools-within-schools; as well as personalization strategies, such as student advisories, family advocate systems, and mentoring programs.

An evaluation by Abt Associates of 119 SLC schools that were funded in the first year of the program (2000) was completed in 2008 (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service, 2008b). The study surveyed all 199 schools in the spring of 2002 and the fall of 2003, and conducted case studies of 18 schools with high student participation and SLC implementation. The study found that SLC schools adopted a number of the recommended structures, with the most common being:

- *Career Academies,* a type of school-within-a-school that focuses on one or more careers and that integrates academic and occupation-related classes (55 schools);
- *Freshman (Ninth Grade) Academies,* designed to bridge middle and high school (58 schools).

In addition, schools adopted a number of complementary strategies, the most common being block scheduling, career clusters or pathways, teacher teams, and adult advocates or mentors.

The study identified a number of factors that facilitated the implementation of the SLCs: strong school leadership, involved and supportive districts, sufficient professional development, sufficient resources (space and instructional materials), and a variety of teacher variables (attitudes toward reform, pedagogical practices, and expertise). The study also identified a number of factors that impeded implementation: staff and administrator turnover, prescriptive district oversight, and scheduling and logistic issues (p. 11).

The study examined a variety of student outcomes, although the study design was limited by a lack of valid comparison groups and by data limited to the second year of implementation. Nonetheless, comparing data after receiving the grants with data from the prior four years showed overall only modest or neutral changes in student outcomes, with a good deal of variation among schools (p. 13). More specifically, the SLC schools showed improved levels of student involvement in extracurricular activities, significant increases in 9th to 10th grade promotion rates, and a downward trend in violence. But there were no significant trends in

academic achievement as measured by either scores on statewide assessments or college entrance exams (p. 15).

Striving Readers Program

The final program, the Striving Readers Program, focuses on instructional improvement. The program aims to (1) raise middle and high school studentsô literacy levels in Title I-eligible schools with significant numbers of students reading below grade levels, and (2) build a strong scientific research base for strategies that improve adolescent literacy instruction.[19] The program, which began in 2005, funded projects in eight school districts. Each project includes a supplemental literacy intervention that targets students who are reading significantly below grade level, a schoolwide program to improve literacy for all students, and a strong evaluation component. To date, only first-year evaluations have been completed that focus on the initial implementation of the projects.

The Role of State Governments

While states have long been concerned with improving their education systems, the growth of state and federal accountability has leant more urgency to their efforts. States have taken a particularly strong, collective interest in the issue of high school reform in the last few years. In 2005, the National Governorôs Association and Achieve sponsored the National Education Summit on High Schools and issued a report, *An Action Agenda to Improve America's High Schools*, that made five recommendations (p. 5):

- **Restore value to the high school diploma** by revising academic standards, upgrading curricula and coursework, and developing assessments that align with the expectations of college and the workplace.

[19] Information retrieved July 17, 2009, from: http://www.ed.gov/programs/strivingreaders/index.html

- **Redesign the American high school** to provide all students with the higher-level knowledge and skills, educational options, and support they must have to succeed.
- **Give high school students the excellent teachers and principals** they need by ensuring teachers and principals have the necessary knowledge and skills and by offering incentives to attract and retain the best and brightest to the neediest schools and subjects.
- **Hold high schools and colleges accountable** for student success by setting meaningful benchmarks, intervening in low-performing schools, and demanding increased accountability of postsecondary institutions.
- **Streamline educational governance** so that the K-12 and postsecondary systems work more closely together.

In January 2009, four state-level organizationsð National Governors Association Center for Best Practices, National Conference of State Legislatures, National Association of State Boards of Education, and Council of Chief State School Officers, issued a joint report, *Accelerating the Agenda: Actions to Improve America's High Schools*, that renews the agenda, documents progress to date, and discusses the remaining challenges in moving forward.

In the agenda item to redesign high schools, the new report provides more emphasis on ñexpanding the supply of high-quality schools through new models such as early college high schools and alternative delivery mechanisms such as charter schools and virtual schools; and preventing students from dropping out and reengaging out-of-school youth through youth development programs and alternative high schoolsò (p. 2), while acknowledging the need to transform or close existing, low-performing high schools ñby seeding innovative practices, evaluating progress, sustaining promising efforts, replicating effective models, and closing persistently failing high schoolsò (p. 13). The agenda item also restates the need to ñidentify and support struggling studentsò and to ñreengage out-of-school youthò (pp. 14-17). These recommendations cover several of the general reform strategies discussed earlier: student support programs, alternative programs, comprehensive school reform, and new schools. What is not

specified is the specific nature of these redesign strategies or how they can be most effectively carried out.

Many state organizations have initiated large-scale efforts to reform their high schools, but few evaluations have been conducted to date.[20] In addition, state governments have initiated school reform efforts as part of their accountability systems. Because the focus of this paper is on efforts that have been evaluated, we highlight one such effort in California.

California established a state accountability system in 1999 known as the Public Schools Accountability Act (PSAA). The PSAA created a system of rewards and sanctions for meeting or not meeting specific performance targets, and established assistance programs for low-performing schools. In 2001, the High Priority Schools Grant Program (HPSGP) was established as part of PSAA to provide additional funds to the lowest-performing schools in the state, taking the place of an earlier program. Participating schools were required to develop an Action Plan (or use one previously developed) to serve as a blueprint for the school and community to focus on improving student achievement and meeting growth targets, and they received $400 per student per year for three years (and a possible fourth year, depending on progress) to use toward implementing improvement strategies. The program allocated almost $1 billion to under-performing California schools.

In 2005, the California Department of Education (CDE) contracted with the American Institutes for Research (AIR) to examine the implementation, impact, costs, and benefits of the HPSGP (Harr, Parrish, Socias, & Gubbins, 2007). The evaluation found that while participating schools demonstrated academic gains over the period of program implementation, their gains

[20] See, for example, the Colorado Childrenôs Campaign (http://www.coloradokids.org/) the Ohio High School Transformation Initiative (http://www.kwfdn.org/), and the Texas High School Project (http://www.thsp.org/).

were not statistically different from the gains of the comparison schools (p. 3). The findings were similar to evaluations of the earlier school improvement program. The study concludes:

> ...it appears that a short-term categorical approach to school reform is *insufficient* to overcome much larger system inadequacies that fail to provide the kinds of longer term support and assistance needed to substantially and consistently improve student performance in the stateôs most challenged schools. We suggest terminating categorical interventions like the HPSGP in favor of more comprehensive statewide school reform that provides long-term administrative and resource support to the stateôs lowest-performing schools enrolling our most academically challenging students (p. 4).

Private Efforts

A number of large-scale school reform efforts have been mounted outside federal and state governments, some focusing on high schools and some focusing on all levels of public schools.

New American Schools

One of the most ambitious efforts to reform American schools was the New American Schools (NAS) program, which ñwas formed in 1991 to create and develop whole-school designs that would be adopted by schools throughout the country in order to improve student performanceò (Berends, Bodilly, & Kirby, 2002, p. xv). NAS was established as a non-profit organization and was largely funded by the private sector. The founders shared the premise underlying other CSR effortsð that piecemeal, programmatic approaches were insufficient to improve low-performing schools. Instead, the premise underlying NAS was:

> All high-quality schools possess, de facto, a unifying design that allows all staff to function to the best of their abilities and that integrates research-based practices into a *coherent* and *mutually reinforcing* set of effective approaches to teaching and learning for the entire school. The best way to ensure that lower-performing schools adopted successful designs was to fund design teams to develop ñbreak the moldò school designs that could be readily adopted by communities around the nation. After developing the design, teams would go on to implement their designs in schools throughout the country. This adoption

> would lead to NASôs primary goal of improving the performance of students (Berends et al., 2002, p. xv).

The NAS approach was systematic and comprehensive, and was carried out in four phases:

1. Competition and selection (1992);
2. Development phase of one year (1992ï 1993);
3. Demonstration phase of two years (1993ï 1995); and
4. Scale-up phase of three years (1995ï 1998).

RAND undertook a long-term evaluation of the NAS effort, documenting the implementation of NAS and the conditions that promoted or impeded the goal of widespread adoption and use of the NAS models as a means to improve student achievement. In particular, the RAND evaluation documented an evolving *theory of action* underlying the NAS approach, from one that schools would simply adopt effective designs in some unspecified way, to one where design teams would need to provide assistance to schools to successfully implement the designs, to one that recognized the importance of school- and district-level factors. One important finding from the evaluation was: ñThe causal chain of events leading to strong implementation and outcomes have proven to be more complex than originally considered by NAS and one that remained largely outside of its control and influenceé in keeping with the literature on implementation indicating the complexity of the change processò (Berends et al., 2002), p. 147). Specifically, RAND identified a number of influential factors:

- The design itself and its ability to offer coherent, comprehensive, and consistent education programs as well as assistance offered by the design teams to schools to ensure implementation;
- The efficacy of the selection and matching process between designs and schools to ensure teacher ñbuy-inò to the design;

- The capacity of the specific schools for undertaking the reform, including the schoolsô other efforts at reform, educational leadership, and teaching capability;
- School-specific demographics, structure, and climate;
- District contexts, including the existing infrastructure supports and incentives for design implementation and improved student performance; and
- Other factors, such as state contexts of testing and accountability, community contexts, and NAS funding policies.

Another important finding was that implementation varied widely by jurisdiction, by design, and across schools (p. xxxii), consistent with other evaluations reported earlier and with the research literature (Desimone, 2002; McLaughlin, 1990). These two findings help explain another major finding: the initial hypothesis (that by adopting a whole-school design a school could improve its performance) was largely unproven (Berends et al., 2002, p. xxxvi)**.** The RAND evaluation goes on to suggest that ñinterventions need to address systemic issues that can hinder implementationò (p. 149).

<u>The Annenberg Challenge</u>

In December 1993 the Annenberg Foundation donated $500 million to public education. The money was used to initiate reform efforts in 2,400 schools in 300 districts in 25 states through a series of 18 school improvement projects in a number of large cities and regions called ñChallenge sites,ò including the San Francisco Bay area, Los Angeles, Houston, South Florida, Chicago, Philadelphia, and Boston, as well as in rural America. The Challenge sites raised an additional $600 million in private and public matching funds, bringing the total financial support to $1.1 billion. A final report published in 2002 by the Annenberg Foundation and the Annenberg Institute for School Reform at Brown University, which was funded to serve as a

ñgreenhouse of ideas and school improvement strategies,ò described the philosophy behind the gift:

> The Challenge did not prescribe one strategy to remedy schoolsô shortcomings. It encouraged districts and those working with them to try different approaches in hopes of discovering different pathways to success. The Challenge came with no ideology other than nonpartisanship, but it embraced the unshakable premise that all children can learn and that good teaching was vital to their success. We sought to build and strengthen parent involvement in their childrenôs classrooms, and we looked for allies in the wider community, for we knew that you cannot turn failing schools around without strong community support (Connell, Martin, & Moore, 2002, p. 13).

Although the grants have now ended, some of the Challenge sites have transformed into new organizations that continue to work on school reform efforts.

Although no formal evaluation of the entire Challenge was conducted, Annenbergôs final report provides nine ñlessonsò:

1. Every child benefits from high expectations and standards.

2. Even large gifts like ours are no substitute for adequate, equitable, and reliable funding. Some Challenge members reported that they could have accomplished more by concentrating their efforts.

3. Schools are too isolated. Reaching out to other schoolsð forming networks for mutual support and criticismð can help overcome problems. But there are limits on the impact of the intermediate organizations, such as lacking a direct authority over school personnel and decision making.

4. Schools need lots of allies to do this work. Parents, businesses and foundations all can play a vital roleð but their support must be built.

5. Professional development holds the key to better schools. We found teachersð new ones and veteransð eager to become better instructors, and we helped them do it.

6. We helped students and teachers to get to know each other better. We found ways to make big schools small and small schools better. This reflected the focus on making ñschools smaller, friendlier places, especially high schools,ò although as one participant noted, ñsmaller is better, but itôs not sufficientò (p. 43).

7. Schools need strong leadership, not just from principals and superintendents, but also in the classroom, on school boards, in the community and in state capitals.

8. Schools cannot improve without accountability. However, those who set the policy and allocate resources should also be held accountable.

9. Public education in America is better than its image. Public schools and those who work with them must do a better job of telling their stories.

The report goes on to state that the Challenge was able to: improve the instructional capacity of teachers and schools, engage with outside partners and demonstrate the value of intermediary organizations; and build strong coalitions with businesses, foundations, universities, and grassroots organizations to bolster public will and support for schools (p. 57).

In-depth evaluations were conducted of reform initiatives in some Challenge sites. One was the Bay Area School Reform Collaborative (BASRC) that worked to improve student achievement in public schools located in the San Francisco area. An evaluation was conducted on one specific BASRC initiative that was implemented in five area districts and two to four selected elementary schools within the districts during a three-year period from 2002-03 to 2004-05 (Porter & Snipes, 2006). The initiative was a process-oriented reform strategy, the ñfocus strategy,ò that had three core features (p. ES-2):

- Coaching of superintendents, district and school leaders, and teachers
- Evidence-based decision-making at all levels of the system (for instance, using student achievement data to inform decisions about policy and practice)
- Networks and collaboration among administrators and teachers and within and across districts and schools

The evaluation produced several key findings (p. ES-2):

- In the districts that participated in the focal strategy, there were improvements in achievement during the years of the initiative, but these improvements were either similar to or only slightly greater than improvements in similar districts in the Bay Area that were not part of the focal strategy.
- The evidence suggests that the BASRC focal strategy was not associated with substantial districtwide improvements in average elementary student achievement on state standards reading tests.

- Although BASRC originally intended to support both districts and schools in a ñbottom-up and top-downò approach, in practice the focal strategy primarily served *district* leadership, was not sustained at the school level, and, thus, typically did not lead to specific changes in instruction or specific instructional supports at the school and classroom levels.

The evaluation also noted that the reform efforts were terminated prematurely in three of the five districts and that by the second year the school-level focus faded. The evaluation concludes with the following recommendation:

> Systemic reforms such as BASRC can take a long time to take root. If they do, and if they are translated into effective instructional improvements, the changes in teaching and learning could be substantial and more sustainable than those evoked by other reforms. This evaluation suggests that external support organizations like BASRC might be more successful in helping districts achieve successful systemic reforms if they can simultaneously support both district offices and schools, if they can move beyond building capacity, and if they can help district offices and schools identify and implement effective instructional changes (p. ES-8).

An evaluation was also conducted of another Challenge effort, the Los Angeles Annenberg Metropolitan Project (LAAMP). The focus of this effort was the creation of a new educational structure, known as a School Family, that brought together teachers, administrators, and parents from high schools, and their feeder middle and elementary schools, to promote consistency and stability both within and across schools, and to focus and accelerate existing reform efforts (Herman & Baker, 2003, p. 2). The initiative funded 28 School Families encompassing 247 schools within 15 of the 81 districts in Los Angeles County. Families were charged with developing their own ñlearning plansò to guide teachersô classroom work built upon reform efforts already in place and guided by a set of research-based ñaction principlesò (p. 4). The evaluation documented a number of achievements, including (p. vi):

- Creation of the School Family concept, which in many cases was responsible for productive changes that could not have been realized by a single school working alone.

- Strengthening of schoolsô acceptance of accountability, their focus on performance, and their capacity for self-evaluation, especially in regard to accessing and using student-achievement data.

- Creation of valuable teacher professional development activities and access to new instructional programs, which were especially helpful for the many new and uncredentialed teachers who were hired to fulfill class-size reduction requirements.

- Encouragement of parental involvement in the schools and in childrenôs learning at home, which had demonstrable effects on student performance.

- Demonstration of the potential of stable learning communities for curing many of the ills facing urban schools.

However, the evaluation found no statistically significant difference in student achievement between LAAMP and non-LAAMP schools over the three-year period between 1997-98 and 2000-01; nor was there any indication that LAAMP had a wide impact on classroom practices.

The evaluation identified a number of challenges that influenced this reform effort, including (pp. 33-38):

- Turbulence created by drastic changes in state requirements, such as the introduction of class-size reduction and the abolition of bilingual education

- Competing demands and time pressures

- The need for new ñgroup processò skills to work in the new structure

- Structural changes that did not quickly result in instructional improvement

- Great variation in the implementation and effectiveness of reforms across School Families.

<u>Gates Foundation High School Grants Initiative</u>

Perhaps the most ambitious effort to reform Americaôs high schools is being conducted by the Bill & Melinda Gates Foundation. The initiative is driven by the view that the U.S. system of public education is fundamentally obsolete, as stated emphatically by Bill Gates at the 2005 National Education Summit on High Schools:

> Americaôs high schools are obsolete. By obsolete, I donôt just mean that our high schools are broken, flawed, and under-fundedð though a case could be made for every one of those points. By obsolete, I mean that our high schoolsð even when theyôre working exactly as designedð cannot teach our kids what they need to know today . . .This isnôt an accident or a flaw in the system; it is the system (Gates, 2005).

The basic strategy was to provide grants to intermediary organizations that would start new high schools or redesign existing high schools.

In 2001 the Gates Foundation funded a five-year evaluation of their initiative that was conducted by the American Institutes for Research and SRI International (Evan et al., 2006). The evaluation, which was completed in 2006, examined the activities of 22 grantee organizations, four model schools that were replicated in a number of locales, and dozens of individual schools collecting both qualitative and quantitative data. Like the evaluation of the NAS initiative, the evaluation of the Gates initiative documented an evolving theory of change, beginning with one that focused simply on creating not only small high schools (with no more than 100 students per grade level), but also schools with a number of essential attributes:

- Common focus by students and teachers on a few important goals;
- High expectations to complete a rigorous course of study to prepare for college, career, and citizenship;
- Personalized relationships between students and teachers;
- Respect and responsibility;
- Time for staff to collaborate;
- Performance-based advancement with adequate student support; and,
- Technology as a tool (p. 3).

Over time, the theory changed as a result of experience with initial grantees, external conditions, and the growth of Foundation staff and capacity. The new theory emphasized more clearly specified models coupled with extensive support, more attention to curriculum and instruction,

and more attention to results-oriented models, regardless of the specific attributes found in the model. The theory was further refined to recognize the importance of working with both school districts and state policy. Finally, the theory emphasized the notion of variety and choice of approaches operated by both districts and outside providers so that families could select schools based on their studentsô interests and needs.

The evaluation examined student outcomes, instructional quality, and district-level efforts to provide school choices to at least 30 percent of students. Student outcomes in new schools and redesigned schools were compared to district averages. The evaluation found that student achievement levels in both new and redesigned schools were below district averages, although data from a more limited sample showed that more than half of both types of schools made above-average gains in student proficiency levels in English language arts, and somewhat smaller gains in math. As the authors point out, students in both types of foundation-sponsored schools had eighth grade test scores well below the district average, which means students were often several grade levels behind (pp. 27-28). The evaluation also found that attendance, engagement, and ninth-to-tenth grade promotion rates in new schools were higher than district averages, but not higher in redesigned schools. As the authors further point out, students in new schools were self-selected and therefore may have been more motivated to succeed than students in redesigned schools (p. 26).

With respect to instructional quality, the evaluation found that the rigor and relevance of coursework in both English language arts and math was higher in foundation-sponsored new high schools than in classroom assignments in comprehensive high schools in the same or nearby districts with similar populations, although the levels of rigor and relevance of the math assignments were generally low in all schools (p. 38).

Finally, the evaluation found that in the 17 districts where the foundation worked, the average percentage of students attending schools of choice increased from six percent in 2001-02 to 33 percent in 2004-05 (p. 61). Districts provided locally developed (four districts), externally provided (four districts) or both types of school models (eight districts) (p. 67).

The authors identify several limitations of the evaluation. One is that the schools included in the evaluation were selected during the initial stage of the foundationôs theory of change that focused on models with particular attributes rather than models with proven records (p. 77). Second, the schools had only been in operation for a few years, so the evaluation was unable to examine longer-term impacts that other research indicates is often necessary to show meaningful results. Despite these limitations, the evaluation yielded some useful insights into the challenges facing school reform:

- New schools faced the challenge of recruiting qualified faculty, especially in mathematics, and in retaining teachers because of stressful working conditions;
- Redesigned schools were less successful than new schools in achieving positive student outcomes because it was harder to change the culture and because redesigned schools often focus more on structural than instructional changes;
- Some redesigned schools also created small learning communities that were stratified by academic ability, creating a form of tracking; and,
- Grantee organizations varied in their capacity to support their schools, with secondary curriculum being a common weakness (pp. 78-79).

New Futures

While school-based reform initiatives have produced little widespread impact on high school performance, so too have initiatives focused on collaborative reform strategies. One ambitious initiative was the New Futures Initiative, promoted and funded by the Annie E. Casey foundation beginning in 1988. New Futures was an attempt to build new collaborative structures among

existing public and private institutions in five cities (Dayton, Ohio; Lawrence, Massachusetts; Little Rock, Arkansas; Pittsburgh, Pennsylvania; and Savannah, Georgia) to address the problems of at-risk youth, including dropping out of school. The key strategy was to establish an oversight collaborative in each city with representation from public and private sector agencies to ñidentify youth problems, develop strategies, and set timelines for addressing these problems, coordinate joint agency activities, and restructure educational and social servicesò (White & Wehlage, 1995, p. 24). The collaboratives also included case managers who 1) brokered services among the disparate agencies serving at-risk youth and their families; 2) served as advocates for at-risk youth; and 3) served as the ñeyes and earsò of the collaboratives by providing information and feedback to the group about what reforms were needed.

Evaluations of this ambitious, systemic reform effort found that it did little to reduce dropout rates and other problems of at-risk youth (Wehlage, Smith, & Lipman, 1992; White & Wehlage, 1995). The evaluations found several generic problems in trying to establish community collaboration:

1. **Slippage between policy and action** because case mangers were generally unsuccessful in overcoming the ñturf battlesò among existing agencies and in getting collaboratives to address them;

2. **Discord over reform policies** because of fundamental disagreements over the definitions, causes, and remedies to problems;

3. **Disjuncture between policy and community conditions** because of the top-down organization of the collaboratives that resulted in an incomplete understanding of the problems and hence ineffective policies.

These prob1ems, clearly evident in New Futures school reforms, paralleled those found in the earlier evaluation of restructured schools. In particular, ñmost educators in New Futures schools believed that the problems that created at-risk students were problems inside the students, not inside the school and its curriculumò (Wehlage, Smith, & Lipman, 1992, p. 73).

Hence, as found in the other systemic reform efforts, there was little incentive or support for changing the fundamental functioning of schools.

Collaborating for Education Reform Initiative

The Collaborating for Education Reform Initiative (CERI) was a new effort at school improvement initiated by the Ford Foundation in 1997. The effort funded eight sites around the U.S. and in Puerto Rico to create collaboratives of community-based organizations to initiate education reforms in their local areas:

> This strategy for school improvement emphasized changing the organization and culture of schools, their relationships with their stakeholders, and the systems in which they were embedded. The Ford sponsors believed that the more coherent, steady, and coordinated these multiple approaches were, the more likely they were to succeed where other disjointed or discontinuous efforts had failed (Bodilly, Chun, Ikemoto, & Stockly, 2004, p. xv).

The collaboratives were directed to work initially with a cluster of schools within districtsð usually a high school and its feeder middle and elementary schoolsð although the goal was to create systemic changes across the entire district (p. xvii). The number of collaborative members varied from five in one site to 19 in another and included local community-based organizations, local colleges and universities, educators, parents, and concerned citizens.

RAND conducted a formative evaluation of the effort from the fall of 1999 to the spring of 2003 that monitored the progress of the collaboratives during their first four years of operation and identified the lessons learned or promising practices from the effort (Bodilly et al., 2004). The study evaluated the collaboratives along five dimensions:

1. **Networks**. All the grantees developed inter-organizational linkages, although some sites made stronger progress than others.
2. **Instructional Improvement**. Four of the grantees contributed to the development and implementation of high quality teaching and support of their partner schools, although they were never fully implemented.

3. **Policy**. Three of the grantees made progress toward changing policy.

4. **Sustainability**. Two of the grantees made significant progress toward sustaining the collaborative and its focus.

5. **Student Achievement**. Very little improvement in student achievement attributed to the collaboratives was evident across any of the sites (pp. xix-xx).

The evaluation identified several factors that accounted for the differences in progress across the four sites, most related to the difficulties in establishing the collaboratives themselves, such as the amount of time and member interaction necessary to build levels of trust for collaboratives to function (p. xxi). Other factors were:

- Inclusion of stakeholders integral to the local context and able to contribute to the collaborativeôs goals
- The perceived legitimacy and authority of the lead organization
- How collaborative members worked together
- The characteristics of and action by the collaborative leadership
- The fostering of the collaborativeôs legitimacy and reputation over time
- The matching of goals to the local context
- The adept use of data to inform theories of action and activities; the habit of continuously reflecting on work and the use of data to alter strategies as necessary
- Early attention to a plan for institutionalizing systemic change, including strategies for sustaining the collaborative as well as sustaining and scaling-up the reform agenda

<u>Schools for a New Society</u>

In 2000, the Carnegie Corporation of New York started a major reform initiative focused on redesigning high schools in urban communities. The initiative, *Schools for a New Society*, provided grants of $60 million (with support from the Bill & Melinda Gates Foundation) to seven citiesð Boston, Chattanooga, Providence, Sacramento, San Diego, Worcester, and

Houstonð to undertake districtwide reform (Carnegie Corporation of New York, 2004). The initiative has several critical components (p. 3):

- Encouraging and supporting partnerships between businesses, universities, parent and student groups and community organizations committed to high school reinvention
- Holding all schools accountable for helping *every* student to meet high standards and to be prepared for participation in higher education, in the workforce, and in confronting the challenges and opportunities of 21st century society
- Raising graduation requirements to ensure that all students take rigorous courses and succeed in them so they are prepared to accomplish their goals in college, the workforce and life
- Transforming large, impersonal high schools into small learning communities or small schools to personalize the student learning experience
- Improving teaching by providing intensive professional development and giving teachers time to work as teams to help all students succeed

The initiative formed a technical support team that not only worked with the seven sites to implement the change framework, but also created a learning community composed of the support team, the foundation, and the sites themselves. This learning community released a report that describes an emerging and deeper framework for redesigning local education systems that includes the following elements:

- A portfolio of schools that offers a range of high-quality options to meet the diverse needs of all students
- A redesigned district to support the new array of schools and the teachers and leaders who work in them
- Engaged youth who are active in their own learning and also contribute to and support the redesign
- An engaged community to demand and support the redesign
- A working partnership with other organizations to enhance capacity and support sustainability

To date, no summative evaluation of the impact of this initiative has been completed, although an evaluation in one site found uneven impacts on reform efforts and student achievement (Reyes, Alexander, Fuller, & Phillips, 2007).

New Century High Schools

The Carnegie Corporation helped to support a separate high school reform initiative in New York City, known as the New Century High Schools (NCHS) initiative. The initiative, with additional support from the Bill & Melinda Gates Foundation and the Open Society Initiative, provided $30 million to start 75 new high schools between September 2002 and September 2005 with leadership and technical support provided by an intermediary organization, New Visions for Public Schools. Funding was provided to private non-profit organizations working in conjunction with individual schools. The intervention theory was based on ten principles:

- *Rigorous instructional program,* enabling every student to master challenging skills, content knowledge, and state standards through relevant, individualized, in-depth, and inquiry-based teaching
- *Personalized relationships* between students and teachers, characterized by close continuous communication and each student having at least one adult to coordinate the support needed for the student to achieve educational goals
- *Clear focus* on teaching and learning and corresponding expectations that every student will succeed
- *Instructional leadership* through effective collaboration and schoolwide support for teaching and learning
- *School-based teacher-driven professional development and collaboration* that is results-driven, standards-based, and embedded in the daily work of the school
- *Meaningful continuous assessment* to diagnose student needs and improve instruction
- *Community partners* that offer support and opportunities for students, families, and the school community and contribute significantly to the schoolôs planning processes, governance, and operations

- *Family/caregiver partnership and involvement* in governance and the design of the schoolôs education program
- *Youth participation and development*, characterized by student voice in teaching and learning and an educational focus on the development of studentsôsocial and emotional skills
- *Effective uses of technology and other resources,* including print, visual, audio, and electronic resources to facilitate learning and school operations

An evaluation of the initiative in 2005-06 found that students in 10 of the initial group of 12 NCHS high schools had graduation rates 18 percentage points higher than students in comparable schools, although the percentage of students earning state-endorsed (Regents) diplomas were comparable (Foley, Klinge, & Reisner, 2008). Students in subsequent entering classes of NCHS had lower rates of school attendance and credit accrual, and higher rates of school suspension, which may impact subsequent graduation rates. The evaluation found that the most important school-level predictor of student performance was what the evaluators labeled ñthe quality of instructional systems,ò a construct that included measures of the ñperceived alignment of instruction with Regents standards, agreement on educational focus, the effectiveness of principal leadership the quality and amount of professional development, teacher influence, and professional collaboration on instructionò (p. ii). The evaluation also found considerable variation in these qualities among the schools.

What Was Learned

Two major lessons can be drawn from these large-scale reform efforts, which can be compared to some lessons drawn from a review of a number of federal reforms efforts of the 1960s, ô70s, and ô80s conducted by RAND researchers (Elmore & McLaughlin, 1988; McLaughlin, 1987).

1. **Large-scale reform initiatives have generally not been successful in making widespread, significant improvements in classroom teaching or in student outcomes**.

The reform initiatives reviewed above supported a number of different, largely schoolwide reform strategies: school restructuring, comprehensive school reform, new schools, and school-community collaborations. No single reform initiative had an overall impact on classroom teaching or on student outcomes, including improved academic achievement and graduation rates, supporting the conclusion stated earlier that no single reform strategy is inherently better than the others. A similar conclusion was reached in a review of earlier federal programs by RAND researchers, who found that:

> Educational reform has historically had little effect on teaching and learning in classrooms. In this pessimistic sense, educational reform is ñsteady work.ò That is, the rewards are puny, measured by substantial changes in what is taught and how; but the work is steady, because there is a limitless supply of new ideas for how schools should be changed and no shortage of political and social pressure to force those ideas onto the political agenda (Elmore & McLaughlin, 1988, p. 3).

2. **There was widespread variability in both the implementation and impact of the initiatives across the settings of schools, districts, and states**.

All of the evaluations found substantial variability in the implementation and impact of the reforms among schools, districts, and states. For instance, the review of the federal Comprehensive School Reform Program found widespread variability in the number of the 11 components of CSR models that schools adopted, with most schools adopting relatively few of them (Aladjem et al., 2006). The RAND evaluation reached the same conclusion: ñvariability is the rule and uniformity is the exceptionò (p. 34), although it also found that ñreforms succeed to the degree that they adapt to and capitalize upon variabilityò (Elmore & McLaughlin, 1988, p. 35).

Although the evaluations of these large-scale school improvement efforts found little impact on either instructional practice or student achievement, they did identify a number of factors that limited their implementation and impact. These factors have also been identified in the growing research literature on the implementation of programs and policies.

Will and Capacity

The research literature has identified two broad factors that affect implementation: *will* and *capacity* (McLaughlin, 1987; McLaughlin, 1990). *Will* and *capacity* refer to traits of both individuals and institutions. At the individual level, *will* refers to the motivation and commitment of educatorsð teachers and administratorsð to implement reform strategies. Both groups of educators play important, yet different, roles in the reform process, and respond to different incentives. Teachersô willingness to engage in reform, which often involves learning new instructional practices, depends on several factors: (1) whether they believe it will benefit their students; (2) whether they believe they have the ability and support to learn new practices; and (3) whether they believe that the requisite effort is professionally rewarding and a district priority (Elmore & McLaughlin, 1988, pp. 42-44). Principalsô willingness to engage in reform depends on similar factors: (1) how they perceive the reform to effect the administrative business of the schoolð whether it may create undue burden to the overall functioning of the school, for example; (2) whether they, too, have the ability and support to implement the reform; and (3) whether they also believe it is professionally rewarding and a district priority (Elmore & McLaughlin, 1988, pp. 45-47).

The individual *capacity* of teachers and administrators to carry out reforms is clearly important. The capacity of teachers to implement reforms, which, again, usually means changing their instructional practices, is a time-consuming, multi-stage process that includes persuasion over the need for reform. Some scholars refer to this as ñsensemakingò (Louis, Febey, & Schroeder, 2005)ð an understanding of the new practice, along with feedback about the

development of new roles and skills, time, collegial interaction, and technical assistance (Elmore & McLaughlin, 1988, pp. 45-47).

While the will and capacity of individuals are clearly important in educational reform, equally important are the will and capacity of institutionsð particularly schoolsð where most educators work. Institutional will resides in the incentives, structures, norms, and culture of the organization that influence the will of individual educators to engage in reform efforts:

> é if you ask teachers to change the way they deal with students and to relate to their colleagues differently, the incentives that operate at the organizational level have to reinforce and promote those behaviors (Elmore, 2004b, p. 38).

The *culture* (or climate) of the school is reflected by a number of identifiable characteristics related to both a schoolôs organization and structure (e.g., school site management, parent involvement and support, maximized learning time), and to processes (e.g., collaborative planning and collegial relationships, sense of community, clear goals and high expectations). The particular ñmixò of factors varies from school to school and thus must develop over time from within rather than through external fiat (National Research Council, Committee on Increasing High School StudentsôEngagement and Motivation to Learn, 2004; Purkey & Smith, 1983; Sarason, 1996).

A number of scholars have concluded that a key factor in reforming a school and changing its culture is the institutional capacity of the school (Malen & Rice, 2004; Newmann, King, & Rigdon, 1997; Roderick, Easton, & Sebring, 2009; Newmann, King, & Youngs, 2000). School capacity consists of several components, with resources being one of the most important. Several types of resources contribute to school capacity. One is *material resources*ð teachers, textbooks, facilities, etc. Another is *human resources*, which represents the skills and abilities of teachers, administrators, and staff. Still another is referred to as *social resources*, which represent the social relationships or ties among students, parents, teachers, and administrators. One in-

depth study of school reform in Chicago found that one particular social resource necessary for school improvement is *relational trust*, which represents the reciprocal, social exchanges among all the participants in the schooling enterprise that depend on respect, competence, personal regard for others, and integrity:

> We view the need to develop relational trust as an essential complement both to governance efforts that focus on bringing new incentives to bear on improving practice and to instructional reforms that seek to deepen the technical capacities of school professionals. Absent more supportive social relations among all adults who share responsibility for student development and who remain mutually dependent on each other to achieve success, new policy initiatives are unlikely to produce desired outcomes. Similarly, new technical resources, no matter how sophisticated in design or well supported in implementation, are not likely to be used well, if at all (Bryk & Schneider, 2002, p. 144).

Social resources and relational trust contribute to a schoolôs professional community (and culture), which has been shown to impact teachersô *collective* responsibility for student learning, and their instructional practices (Lee & Smith, 2001; Little, 2003; McLaughlin & Talbert, 2001; Louis, Marks, & Kruse, 1996; Louis & Marks, 1998). Such collective responsibility contributes to a schoolôs internal accountability, in which school staff set clear standards for student performance, monitor student success, and exert peer pressure to achieve school goals (Elmore, 2004b, Chapter 4; Newmann et al., 1997).

Another important component of school capacity is *productivity*, which is ñthe ability of the school to translate resources into expected outcomesò (Malen & Rice, 2004, p. 635). One common criticism of schools is that they do not use their resources efficiently: ñThe fundamental problem is not a lack of resources but poor application of available resourcesò (Hanushek, 1996, p. 30). Productivity depends, in part, on not just the amount and types of resources and programs (including their standards, curriculum, and assessment), but whether they are aligned with each other and working together to achieve the same objectives (Grubb, 2009; Newmann et al., 2000). Productivity also depends on the coherence between current reform efforts and other reforms and

policies that may be placing additional or conflicting demands on the school that can impede or undermine the current reform efforts (Newmann, Smith, Allensworth, & Bryk, 2001). Additionally, the reform itself may undercut productivity and dilute capacity by the demands it places on the organization, such as the need to adopt new programs and produce quick results in response to accountability demands (Malen & Rice, 2004).[21] The evaluation of the Annenberg challenge grant in Los Angeles, for example, cited the turbulence from changes in state policy as impeding on attempts to improve instructional practice in the schools (Herman & Baker, 2003).

If will and capacity are fundamental to successful school reform, one important question is whether both must be present in sufficient amounts before reform strategies can be successfully implemented, or whether the reform itself can sufficiently alter the will and capacity of educators and schools to ensure successful implementation. A study of school reform in Chicago found that indicators of school capacity predicted whether reforms would work, with high-capacity schools showing three times greater impact than low-capacity schools (Roderick et al., 2009, p. 17). The lack of capacity in teachers, administrators, schools, and districts was cited in a number of the evaluations as the major reason for the weak implementation and impact of several major reform efforts, as well as the widespread variability among schools and sites. Some evidence also suggests that capacity is generally lower in high schools than in middle or elementary schools (Newmann et al., 1997).

Although all reform initiatives include a combination of materials, professional development, and technical assistance to build capacity, in many cases the support was unable to raise capacity sufficiently to ensure successful implementation. Many reform strategies attempt to ensure sufficient will by requiring votes of faculty before initiating reforms, although the process for securing buy-in is sometimes coercive and fails to secure genuine willingness to

[21] Malen and Rice refer to these organizational responses as ñorganizational freneticism and fragmentation.ò

implement the reforms (Aladjem et al., 2006; Datnow, 2000). The study of Chicago school reform found that relationship trust (a component of school capacity) is easier to build in schools that both students and teachers have chosen to attend, thereby assuring a shared commitment (Bryk & Schneider, 2002, p. 142). This literature suggests that it may be necessary to recruit and select teachers with at least some pre-existing level of will and capacity before undertaking school reform. This is one advantage that charter schools typically enjoy over comprehensive school reformð charter schools select their teachers (and may not be required to hire union, district teachers), while comprehensive school reform models are implemented in existing schools and often with existing staffs.

Yet other research suggests that, at least in some instances ñé belief or commitment can follow mandated or coerced involvement at both the individual and system levelò (McLaughlin, 1990), p. 13). For example, changing practice and showing teachers that effective practices work can change their beliefs about studentsô ability to learn (Guskey, 1989). Some reform developers point out an important distinction between the will of district and school leadership versus the will of teachers:

> District leaders remaining on the sidelines, ñwaiting and seeing,ò passively undermining grassroots efforts will trump even strong school staff (administrator and teacher) buy-in every time. Top district leaders and those running key central office functions must be on board from the get-go. Teacher buy-in is important but comes *after* leadership buy-in (district and building) and for many staff only comes after implementation and early successes (Connell, Legters, Klem, & West, 2006, p. 6).

In addition, certain organizational features, such as smaller school size and shared decision-making, may facilitate the development of teachersô will and capacity even as they may require new skills and commitment to effectively operate in the restructured organizational arrangements (Newmann, 1993; Ancess, 2000).

The issue of whether or the extent to which a certain level of will and capacity must exist before reforms are introduced is a difficult one to resolve. On the one hand, greater pre-existing levels of will and capacity should helpð but by no means guaranteeð more successful implementation. On the other hand, low levels of will and capacity can be addressed if reforms are properly designed, funded, and matched to the existing needs and capacity of the school, assuming the pre-existing needs and capacity of the school can be determined.

The evaluations and existing research literature have identified a number of more specific factors that influence the implementation and impact of school reform strategies, at least in part, through their impact on will and capacity.

The Nature of the Reform

Virtually all reforms, no matter what their type, are designed to improve the capacity of educators and schools to improve student outcomes. Several features of the reform affect successful implementation. One is whether the reform is locally or externally developed. The research literature suggests that externally developed reform strategies are generally faster and easier to implement in a local site (Desimone, 2002; Glennan Jr. et al., 2004). Another feature of the reform strategy is its specificityð the degree to which it is prescriptive and content-focused versus more process-oriented. Instructional strategies appear to be more successful if they focus on pedagogy rather than curriculum or particular instructional programs (Slavin et al., 2008; Slavin et al., 2009). In contrast, comprehensive school reform models that emphasize prescriptive approaches seem to be more successfully implemented than those with more process-oriented approaches (Desimone, 2002). The need for specificity may depend on existing capacity. One view is that excellent teachers need more autonomy in their choice and application

of instructional approaches, while weak or novice teachers benefit from more guidance (Porter, 1989). Thus, instructional programs that are uniformly prescriptive may be less effective than flexible ones tailored to the capabilities of the teacher and the needs of the students. A recent evaluation of two supplemental literacy programs for struggling ninth-grade readers appears to bear this out, finding larger effects in the flexible program than the prescriptive one that required teachers to deliver course content in precise and systematic fashion (Corrin, Somers, Kemple, Nelson, & Sepanik, 2009). A final feature of the reform strategy is the amount and type of support provided in the form of materials, time, professional development, and technical assistance from district staff or outside technical assistance providers. Technical assistance includes developing implementation benchmarks, monitoring the implementation process using these benchmarks, and providing feedback (Desimone, 2002; Berends et al., 2002).

Selection Process

The selection process is generally considered a critical factor in building will and support to implement for reform. Research has shown that teacher buy-in is essential for mounting and sustaining reform efforts, with the principal playing a crucial role in promoting teachersô active participation in the selection process (Berends et al., 2002; Desimone, 2002; Glennan Jr. et al., 2004). Some comprehensive school reform models, as part of their design, work toward building teachersô and principalsô understanding of the reform and, in some cases, help shape the reform to the needs and capacity of the school, a process known as ñadaptive implementationò (McLaughlin, 1990, p. 14). Yet at least one study of the federal Comprehensive School Reform program found that the selection process was less important in ensuring successful

implementation than the school-level supports that were provided, supporting the idea that will can be altered after successful implementation (Aladjem et al., 2006).

Context

A number of contextual factors at the school, district, state, and federal levels affect successful implementation. At the school level, studies have found that reforms are easier to adopt in smaller versus larger schools, and in elementary versus secondary schools (Berends et al., 2002; Bryk and Schneider, 2002; Desimone, 2002). Another contextual factor is teacher and administrator stability; it is harder to mount and sustain reform efforts if there is a lot of staff turnover, although features of the reform and the process of selection can affect stability by providing support and promoting staff buy-in. There is no consistent evidence on whether student demographics matterð the NAS evaluation did find implementation levels were lower in schools serving significant populations of both poor and minority students (Berends et al., 2002). At the district level, studies have found that implementation is more successful in districts that are more supportive and have more stable leadership that backs the reform efforts; provide more school-level autonomy; have a more coherent program of reform; are not burdened with crises or other distractions; and have trusting relationships among district staff, the school board, and the unions (Aladjem et al., 2006; Berends et al., 2002; Desimone, 2002). At the state and federal levels, existing policies, particularly those related to meeting short-term accountability requirements, impact implementation by creating program incoherence and organizational inefficiency at the school level.

Beyond simply identifying factors that promote or impede reform, it is important to identify a *theory of action* to explain the underlying causal process of how reform is supposed to

bring about improved student and school performance. A number of prior reform efforts had an explicit theory of action or at least a stated premise underlying the reform initiative. In some cases, as in the Gates initiative, the theory changed over time in response to initial monitoring of the reform effort. But it is probably safe to say that the conclusion drawn from the RAND evaluation of the NAS applies to virtually all reform initiatives: that the theory of action underlying the reforms ñwas largely under-developed and underspecifiedò (Berends et al., 2002), p. 147).

Implications

The evaluations of past reform efforts and the existing research literature on implementation provide a number of implications for developing more successful efforts in the future:

Improvement strategies, especially more comprehensive ones, will not be successful until critical aspects of capacity and context are improved. This was one of the major implications from the NAS evaluation (Berends et al., 2002, p. 147). Most evaluations cited school and district capacity as critical factors in implementation success. The capacity of individualsð teachers and administratorsð as well as the institutional capacity of the school itself are key factors to successful implementation. School capacity depends on having sufficient and correct alignment of resources (including sufficient time); it also depends on coherence in its efforts across all the demands placed on schools and their staffs by districts, as well as state and federal policy requirements. Building capacity also depends on having sufficient will or readiness, especially among school and district leadership, to build capacity and initiate reform. Context factors in the school (such as student composition and the stability of school personnel) and in the district (such as the harmony between the superintendent, school board, and teachersô

union) are much harder to improve.

Districts should be the focal point of high school improvement efforts. Districts play a critical role in initiating and sustaining school improvement efforts (Coburn & Talbert, 2006; Spillane, 1996; Supovitz, 2006). They also can help create coherence and improve productivity in individual schools engaging in reform efforts by coordinating all the reform efforts within the districts. Moreover, districts that take more control in selecting school improvement strategies tend to provide more support for those efforts (Aladjem et al., 2006). Finally, districts can create schools, including alternative education schools, so they are in the best position to create a coherent, systemwide effort that involves reforming existing high schools and creating new schools to meet the needs of all students in the districts.

Successful improvement efforts depend not just on selecting the most appropriate strategy, but also the most appropriate technical assistance provider. Although some schools and districts are able to develop local improvement efforts, most require assistance from outside developers and technical assistance providers for materials, information, professional development, implementation monitoring and guidance (Glennan Jr. et al., 2004). But technical assistance providers vary in their capacity to provide sufficient support to schools and districts. Therefore, it is important that schools and districts select a technical assistance provider that meets their needs and can provide the proper support (Glennan Jr. et al., 2004; Supovitz, 2006). In particular, it is important for technical assistance providers to provide consistent, clear, and frequent communication and assistance to districts and schools (Berends et al., 2002, p. 149).

Improvement strategies need sufficient time to be fully implemented and to observe meaningful changes in student outcomes. Improving high school performance requires making fundamental changes in the beliefs and practices of school personnel, and ultimately in changing

the culture of the school, including the relationships between students and teachers and among school staff (Bryk & Schneider, 2002; Sarason, 1996; Purkey & Smith, 1983). Such changes, especially those concerning instructional practice, take years. Yet because many evaluations tend to focus on the first years of implementation, they may not be sufficient to observe changes in instructional practice that result in improved student outcomes. This implies that evaluations of improvement strategies should be done over long enough periods of time to allow those changes to occurð probably five years or more.

Rethinking the Federal Role in Promoting School Improvement

The federal government has played a significant role in the past in promoting school reform, but with very limited impact. The recent change in administrations, the new federal stimulus funds, and the expected reauthorization of the No Child Left Behind Act together provide an unprecedented opportunity to rethink the federal role in improving high school performance.

The Role of Policy in Reforming Schools

Policymakers have a limited number of mechanisms or policy instruments at their disposal to effect change in public schools, and they often lack an explicit understanding of how the mechanisms they employ are supposed to workð that is, they lack an explicit theory of action (McDonnell & Elmore, 1987; McDonnell, 2004).

The five policy instruments are: (1) mandates (2) inducements (3) capacity-building (4) system-changing (5) and hortatory. Each policy instrument employs specific mechanisms that are suited for particular types of educational problems (see **Table 3**). Underlying all the policy

mechanisms are assumptions about the existing will (behavior) and capacity of policy targetsð individuals and institutionsð and the ability to alter those features. For example, mandates impose rules on policy targets in order to get them to produce desired outcomes under the assumption that the policy targets have the capacity to comply. Inducements provide money to institutions to produce more desired outcomes under the assumption that the capacity exists or can easily be acquired. In contrast, capacity-building instruments provide money for investment under the assumption that the policy targets lack the current capacity to produce the desired outcome. System-changing instruments go one step further by supporting new entrants (e.g., private service providers or private schools) under the assumption that existing institutions, even with more rules and money, are incapable of producing the directed outcomes. Finally, hortatory instruments attempt to alter behavior of individuals and institutions by providing relevant information that appeals to their values to produce the desired outcomes under the assumption that they will act on the information.

Policymakers often utilize more than one policy instrument, especially when they wish to achieve multiple policy goals and when the problems they are addressing are complex. This is often the case in education. State performance-based accountability systems, for example, often use mandates, inducements, capacity-building, and persuasion to bring about school reform.

Policymakersô choice of policy instruments depends upon (1) how they define the educational problem they are trying to solve, and (2) the range of resources they have and constraints they face, including governmental capacity, fiscal resources, policy support and opposition, and available information (McDonnell & Elmore, 1987). These choices and actions are not just based on technical considerations and available information, but are also based on policymakersô values: ñé policymakers also interpret this information using their own-pre-

existing values about how the system actually works and how it is supposed to work é [and they] prefer policy instruments consistent with their own valuesò (McDonnell & Elmore, 1987, p. 145-6). Because of the complexity of both the educational problems policymakers address and the policymaking process, the theory of action underlying a policy may be wrong, or at least incomplete.

The Limits of Performance-based Accountability

School reform efforts, including those focused on high schools, are embedded in state and federal performance-based accountability systems. The federal-level No Child Left Behind Act, for example, requires states to establish systems of support for schools that are identified for improvement. That includes most of the reform strategies identified earlier, such as implementing a new, research-based instructional program, offering supplemental services from a state-approved provider, restructuring the internal organization of the school, and reopening the school as a public charter (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service., 2009, p. 17).

Performance-based accountability systems attempt to improve student and school performance by:

- Developing content standards that specify what all students are expected to know and do;
- Developing assessments to measure whether students are meeting those standards;
- Specifying performance standards that students and schools are expected to reach and a timetable for reaching them;
- Providing a series of rewards and sanctionsð high stakesð for both students and schools to meet the performance standards.

Performance-based accountability systems have their own theory of action on how schools are

supposed to improve:

> Performance-based accountability systems operate on the theory that measuring performance, when coupled with rewards and sanctionsé will cause schools and individuals who work in them, including students, teachers, and administrators, to work harder and perform at higher levels (Elmore, 2004a, p. 277).

But some critics argue that this theory, too, is ñunderspecifiedò because it remains unknown how schools and the individuals in them are supposed to respond to these rewards and sanctions in order to produce the desired outcomes (Elmore, 2004a, p. 278). How schools respond is largely determined by the degree of the alignment between the requirements of the external accountability system and the schoolôs internal accountability mechanisms that reflect the coherence and agreement around individual responsibility and collective expectations for teaching practice and student learning (Elmore, 2004b , Chapter 4; Newmann et al., 1997). Further undermining could result from the competing demands of state and federal accountability systems. This suggests that performance-based accountability systems may undermine the productive capacity of schools, as described earlier. In addition to an underspecified theory of action, performance-based accountability systems suffer from a number of other, technical, shortcomings, including a lack of valid, reliable, and accurate assessments of student and school performance (Furhman & Elmore, 2004).

Performance-based accountability systems rely on high-stakes incentives as the primary mechanism for reforming schools and improving student achievement, rather than providing resources and support to increase capacity. This assumes that the primary impediment to improved student outcomes is a lack of will rather than a lack of capacity, an assumption that some observers question:

> Is it plausible to assume that educators actually know how to substantially improve student performance, but that they are for some obscure reason withholding this knowledge because they have been insufficiently motivated or rewarded by existing incentive structures (Elmore, 2004a, p. 280)?

While past research on policy implementation clearly recognizes the importance of incentives (Hanushek & Jorgenson, 1996), it also clearly demonstrates that students, educators, and schools require resources and support to improve their capacity (Elmore, 2004b; McLaughlin, 1987, 1990). Moreover, improving capacity may also lead to improvements in will (Guskey, 1989).[22]

The Shortcomings of NCLB

NCLB suffers from many of the same shortcomings as all performance-based accountability systems. Yet despite these shortcomings, studies of NCLB implementation have found that educators generally report positive impacts, including improvements in academic performance and instruction as well as a concerted focus on student learning (Stecher et al., 2008, p. xvii)

NCLB suffers from another more general shortcoming facing all federal initiatives to improve public schools: the fragmented educational system in the United States. The constitutional authority for education in the United States resides in states that, in turn, have delegated substantial authority to the more than 15,000 local education agencies. This fragmentation of authority and power makes it difficult for the federal government, which provides about eight percent of the funding for public education, to effect widespread and meaningful change in educational practice at the local level (Cohen & Spillane, 1992).

Fragmentation seriously limits the success of NCLB. Many of the key provisions of the law specify ambitious federal policy goals, but allow states considerable authority in how they respond to those provisions. For example, while NCLB sets a federal policy goal that all students in the United States should be proficient in key academic subjects by the year 2014, states have the authority to set content and achievement standards. Similarly, NCLB mandates that schools

[22] Even with resources and support, however, some teachersô capacity and will may not change, as was discovered in one recent case study of a professional development activity (Fisler & Firestone, 2006).

receiving Title I funds show adequate yearly progress (AYP) toward reaching the performance target for all identifiable subgroups of students (e.g., racial and ethnic minority, poor, English language learner, and disabled students), but states have the authority to determine how AYP is calculated and the size of the identifiable subgroups. As a result, there is considerable unevenness among states in how many schools and which students in those schools are meeting AYP goals (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service, 2009, p. 18). There are also inconsistencies and conflicts between NCLB as a federal accountability system and state accountability systems that have different policy goals and performance measures.[23]

Another shortcoming of NCLB is the mix of policy instruments it employs. NCLB draws on all five policy instruments in order to achieve its ambitious policy goals, but the mix of instruments is problematic. In particular, NCLB is heavy on mandates and short on inducements and capacity building. NCLB has been criticized, for example, because it fails to provide sufficient financial support for its mandates, such as annual testing, resulting in a claim of it being an ñunfunded mandateò (Sunderman & Kim, 2007).[24] More fundamental is the heavy reliance on mandates and inducements, which evoke short-term responses, and less reliance on capacity-building, which relies on long-term investment to improve capacity. The immediate need to meet the short-term mandates of the NCLB performance goals undermines the ability to provide the investment needed to improve the long-term capacity of individuals and institutions to improve schools (Furhman & Elmore, 2004).

[23] California, for example, has its own accountability system that differs substantially from NCLB (see: http://www.cde.ca.gov/ta/ac/ay/documents/aproverview05.pdf).

[24] In reviewing a number of recent statutes, the Government Accounting Office determined that NCLB was not an unfunded mandate because the requirements it imposes are only conditional upon receiving Title I funds (U.S. GAO, 2005, p. 22).

Under NCLB, for example, failure of schools to meet AYP goals for two years invokes a provision that allows students to transfer to another public school in the district; and failure to meet goals for three years invokes a provision that low-income students be offered supplemental educational services. Both actions drain resources and effort away from improving the long-term capacity of low-performing schools. Moreover, the capacity-building provisions of NCLB are limited, focused on improving the individual capacity of teachers through professional development, not on improving the institutional capacity of schools necessary for deep and sustained school improvement. There is also limited support for developing the stateôs capacity to implement all the provisions of the law, including developing a system of support for schools in need of improvement (Sunderman & Orfied, 2008).

NCLB is particularly problematic for fostering high school improvement (Alliance for Excellent Education, 2007). School choice for high school students is limited because the majority of Americaôs school districts only have one high school and, in many urban districts, the majority of schools are under performing, which limits transfer options. And because of funding formulas, only a small percentage of Title I resources are allocated to high school students, which results in little support for high school improvement and exempts most high schools from undertaking improvement efforts.

Recommendations

The federal role in improving high school performance is closely related to its role in improving all levels of educational performance. It is also closely tied to its role in performance-based accountability, particularly through NCLB. Despite spending considerable resources over a long period of time, federal efforts have produced little improvement in high school performance. To

remedy this situation, a number of proposals have been made to strengthen NCLB.[25] The following recommendations focus on actions the federal government could take to have a greater impact on improving high school performance.

1. **Support the development of broader indicators of student outcomes and progress in high school, and include these indicators in the National Assessment of Educational Progress.**

Federal and state accountability systems generally focus on a narrow range of student outcomes, primarily academic outcomes as measured by standardized tests. Not only do many tests do a poor job of measuring state standards; accountability systems tend to measure performance in a limited number of subjects, resulting in reduced attention to other subject areas, especially for students who are not yet proficient in the requisite subjects. More importantly, academic achievement tests ignore other, noncognitive, outcomes that are critical to success in college and careers, and to effective citizenship. The federal government should support the development of broader indicators of student outcomes that include cognitive and noncognitive skills and graduation rates as well as indicators of student progress toward graduation, such as promotion rates. These indicators should be used to judge both student and school performance. They should also be added to the National Assessment of Educational Progress so that it can provide a more comprehensive picture of student performance (Rothstein et al., 2008).[26]

2. **Work with states to develop mechanisms so that schools and districts with the greatest needs and least capacity select the appropriate strategy to meet their needs, and be given sufficient support to develop the capacity to implement and sustain that strategy.**

Existing research has documented widespread disparities in the capacity of schools and districts to mount and sustain effective high school improvement strategies. Yet under NCLB,

[25] See recommendations made by the Center on Education Policy (http://www.cep-dc.org/index.cfm?fuseaction=Page.viewPage&pageId=536&parentID=481) and RAND (http://www.rand.org/education/features/five_key_education_priorities_for_the_obama_administration.html).
[26] Survey items could also be added to NAEP that ask students about their engagement and exposure to effective classroom practice.

schools and districts are afforded considerable autonomy in selecting from a range of reform strategies that may not match either their needs or capacity.[27] Instead, states need to develop mechanisms so that schools and districts with the greatest needs and least capacity select the appropriate strategy to meet their needs and be given sufficient support to develop the capacity to implement and sustain that strategy. The strategy must include the appropriate level of specificityð schools and districts with the least capacity should select more specific interventions that include benchmarks and timetables to monitor and evaluate their progress in implementing the improvement strategies. Funding should also focus on high schools with the greatest needs, where a majority of the students are racial and ethnic minorities and where high school graduation is not the normð so-called ñdropout factoriesò (Balfanz & Legters, 2004).

The federal government should work with states to develop support systems that assure better matching of reform strategies to the needs and capacity of schools and districts in need of improvement. The systems should identify how states would: (1) assess readiness and capacity of schools and districts to undertake reform; (2) identify strategies appropriate for different levels of readiness and capacity; (3) select technical assistance providers that have the appropriate experience and capacity to support schools and districts in implementing selected strategies and a commitment to build local capacity to sustain reform efforts; (4) ensure coherence with existing state and federal policy demands; and (5) monitor improvement efforts and intervene when necessary.

3. Build the capacity of state governments and technical assistance providers.

State governments play a pivotal role in school improvement efforts and serve as the

[27] For example, under NCLB, schools at the most advanced stage of improvement, called ñschools in restructuring,ò get to choose among five restructuring options, with the most popular being ñany other major restructuring of the schoolôs governance arrangement that makes fundamental reforms, such as significant changes in the schoolôs staffing and governanceò (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service., 2009, p. 29; Mathis, 2009).

major conduit for federal initiatives (Smith & OõDay, 1991). States have major responsibilities under NCLB, including administering, collecting, and reporting school and district student performance data, as well as developing a support system for schools and districts in need of improvement. Yet NCLB provides little funding to states to support these efforts, with one of two NCLB provisions to support school improvement activities not even funded until 2007 (Sunderman & Kim, 2007; Fagan, 2008). If states are going to play a more central and effective role in supporting school improvement efforts, then they need to build their capacity to engage in those efforts and the federal government should help build that capacity through both funding and more explicit guidance on what that capacity should look like (Minnici & Hill, 2007).

For example, states need to have the personnel, expertise, and information to measure the capacity of schools, districts, and technical assistance providersð all key players in school improvement effortsð so that they can more effectively target school improvement efforts in their states. They also need to have the capacity to monitor school and district improvement efforts after they get underway and evaluate their effectiveness. This will require building more robust state-level, educational data systems, a topic discussed further below. A final component of state capacity is to build a mechanism for continuous system improvement by building the capacity of low-performing school districts and schools.[28] One approach would be for states to target initial improvement efforts to a group of ñlighthouseò districts where improvement strategies are implemented and evaluated, with successful strategies replicated in other schools and districts.[29]

The federal government should also help build the capacity of technical assistance providers. The expected increase in federal stimulus funds, especially those focused on school

[28] For an example of one approach, see Rumberger & Connell (2007).
[29] For an illustration of this strategy applied to the problem of solving the dropout crisis in California, see Rumberger (2008b).

improvement and innovation, could greatly increase the demand for technical assistance. To assist states, districts, and schools in identifying technical assistance providers with the proper expertise to work on high school reform efforts, the federal government should consider developing a system for certifying technical assistance providers. The certification would assess providersô experience and evidence of success in working with high schools, such as whether they have developed and implemented proven programs. If the demand for technical assistance providers exceeds the supply of certified ones, the federal government should ensure support to train additional providers and to create state networks of providers to share expertise and experience.

4. Establish coherence among federal policy initiatives, between federal and state initiatives, and between government and foundation initiatives.

One of the impediments to building capacity is the incoherence among competing policy demands placed on schools and districts by states and the federal government:

> Indeed, the fragmented policy system creates, exacerbates, and prevents the solution of the serious, long-term problems in educational content, pedagogy, and support services that have become endemic to the system (Smith & OôDay, 1991, p. 238).

In particular, short-term demands for annual improvements in student outcomes from state and federal accountability systems can subvert efforts to build capacity for long-term, continuous improvement. The federal government can address this problem by granting waivers and encouraging states to grant waiversð a ñstop the clockò provisionð for schools and districts that are engaged in well-designed, research-based improvement efforts, as long as they demonstrate improvement on so-called ñprogressò indicators. Such indicators would include: improvements in classroom instruction, quality implementation of classroom instruction, reduction in disciplinary referrals, student attendance, teacher attendance and engagement, and school climate. This will require developing a more robust educational data system.

The federal government can also improve coherence in the information resources it provides to states, districts, and schools. One of the major sources of support provided by the federal government is information on all aspects of the school improvement process, including effective practices. But there is little coherence in those efforts and, in some cases, the information is contradictory. For example, the federal government supports two directories on effective educational practicesð the What Works Clearinghouse and the Best Evidence Encyclopediað that sometimes reach conflicting conclusions on whether particular reform models are effective.[30] Other federally supported information sources, such as the five national Content Centers that provide technical assistance to the 16 Regional Comprehensive Centers to support school improvement efforts under NCLB, also provide information on effective practices not necessarily drawn from or in agreement with information from the two federally funded directories.[31]

Finally, the federal government should develop a mechanism for coordinating its own efforts to fund the development, evaluation, and scale-up model programs with those of foundations. Foundations have long been involved in the development of specific school improvement models as well as large-scale initiatives to scale up those models, such as the Annenberg Challenge and the Gates Small Schools Initiative. But these efforts may also result in conflicts or confusion with state and federal policy initiatives that diminish their impact. Better coordination of government and foundation efforts could yield more effective outcomes from both sets of efforts (Glennan Jr. et al., 2004, Chapter 16).

---

[30] For example, the Best Evidence Encyclopediaôs report on middle and high school comprehensive school reform models rates the studies of First Things First conclusive, while the What Works Clearinghouse rates most aspects of the same studies as not meeting standards and consequently does not report the findings.

[31] One of the five Content Centers, The National High School Center, focuses on high schools (http://www.betterhighschools.org/resources/). It published a report synthesizing findings from four ñrigorousò evaluations of four high school reform models, including First Things First (Herlihy & Quint, 2007).

5. **Support the development of state and local data systems to measure more than educational inputs and outputs, assessing district and school readiness and capacity to initiate reform as well as processes and practices that measure progress toward improving student outcomes**

States are developing longitudinal data student systems that can provide more useful information on student achievement that can be used to develop more accurate, ñvalue-addedò measures of school performance (McCaffrey, Lockwood, Koretz, & Hamilton, 2003).[32] But to be useful for guiding improvement efforts in schools and districts, educational data systems need to be expanded and linked to all phases of the school improvement process. That process includes: (1) measuring the existing readiness (will) and capacity of schools and districts to better target improvement strategies and determine the amount and kinds of support needed to implement the strategies and build capacity to sustain them; (2) monitoring the implementation of the improvement strategies that can be compared to benchmarks and timelines, including ñprogressò indicators that provide more useful information on the initial impact of the improvement efforts and facilitate timely mid-course corrections;[33] (3) evaluating the effectiveness and the cost-effectiveness (Levin, 1988) of improvement strategies so that successful ones can be replicated in other schools and districts. An example of a more robust data system can be found in Chicago, where the Consortium of Chicago School Research used teacher surveys to develop an indicator of program coherence, a component of school capacity that helped to predict improved student achievement (Roderick et al., 2009, p. 25)ò.

[32] This work is spearheaded by the Data Quality Campaign (http://dataqualitycampaign.org/).
[33] As noted in the evaluation of the New American Schools effort: ñthe best way to measure whether an intervention is having an effect is to measure variables more closely associated with the interventionsò (Berends et al., 2002, p. 150). This idea was also discussed in earlier recommendations to build indicator systems that included ñprocessò indicators (Shavelson, McDonnell, Oakes, & Carey, 1987).

Conclusions

There is a long history of federal, state, and private efforts to bring about substantial improvements in public schools, including high schools. These efforts have largely failed. But the numerous evaluations of these efforts have identified a number of factors that promote or impede the successful implementation of reform strategies in schools. If new efforts at reform, including those now being considered by the federal government, can address these factors, then they should be more successful in substantially improving public schools in the future. The current political climate for change, coupled with the federal stimulus funds, provides an ideal time to try a new approach.

The overriding theme from this review is that for high school improvement efforts to be successful, they must address the will and capacity of both individual educators and educational institutions. Instead of focusing on simply developing and evaluating effective reform strategies and providing support to implement them, the federal government and the state governments must also focus on measuring the will and capacity of schools and districts to initiate reform and on matching reform strategies with those conditions. Reform efforts should also focus on building the capacity of state governments and technical assistance providers to provide sufficient support for statewide improvement efforts. Underlying all these efforts should be more robust state education data systems that provide comprehensive and useful data on all aspects of the reform process, from indicators that measure initial will and capacity, to process indicators that measure implementation quality and student progress, and finally to outcome indicators that measure a broader and more useful set of student outcomes.

The federal government can play a pivotal role in supporting this capacity-building

school reform strategy. It can further develop such a strategy in conjunction with its current efforts to improve schools through the American Recovery and Reinvestment Act (ARRA), although it would require some modification of those efforts (U.S. Department of Education 2009). For example, current efforts focus on raising standards and increasing the recruitment and effectiveness of teachers (individual capacity), not on improving the institutional capacity of schools and districts. Current efforts also focus on building and using data systems to support school improvement, but these efforts involve student and teacher data rather than institutional data on schools and districts. Finally, the federal government should revise the accountability provisions of the No Child Left Behind Act to support high school reform strategies that produce measurable, meaningful, and sustainable progress in developing a broad array of valuable student outcomes, rather than encouraging schools and districts to adopt narrow, unproven strategies in to order to meet short-term performance targets in a few academic subjects. The federal government should also support the development and implementation of strategies in earlier grade levelsð such as preschool programs, class size reduction, and middle school reformð that have been shown to improve high school outcomes (Belfield & Levin, 2007).

High schools play a critical role in preparing young people for their futures. Many studentsð particularly racial, ethnic, and linguistic minority studentsð are not well served by these institutions and, as a result, face a bleak future. The United States, too, will face a bleak future if it cannot enable more students to graduate from high school sufficiently prepared for further education, productive jobs, and active citizenship. The federal government can and should play a more active role in supporting systemic reforms of high schools to meet this challenge.

References

ACT (2008). *The forgotten middle:Ensuring that all students are on target for college and career readiness before high school*. Iowa City: Author. Retrieved July 17, 2009, from: http://www.act.org/research/policymakers/reports/ForgottenMiddle.html

Aladjem, D. K., LeFloch, K. C., Zhang, Y., Kurki, A., Boyle, A., Taylor, J. E., Hermann, S., Uekawa, K., Thomsen, K., & Fashola, O. (2006). *Models matter—The final report of the National Longitudinal Evaluation of Comprehensive School Reform*. Washington, D.C.: American Institutes for Research.

Alexander, K.L., Entwisle, D.R., & Kabbini, N.S. (2001). The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record*, *103,* 760-882.

Allensworth, E. & Easton, J. Q. (2005). *The on-track indicator as a predictor of high school graduation*. Chicago: Consortium on Chicago School Research, University of Chicago.

Alliance for Excellent Education (2007). *In need of improvement: NCLB and high schools*. Policy Brief Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.all4ed.org/publication_material/NCLB

American Diploma Project (2004). *Ready or not: Creating a high school diploma that counts*. Washington, D.C.: Achieve. Retrieved July 17, 2009, from: http://www.achieve.org/ReadyorNot

American Diploma Project (2007). *Aligning high school graduation requirements with the real world: A road map for states*. Washington, D.C.: Achieve. Retrieved July 17, 2009, from: http://www.achieve.org/AligningHighSchoolGradRequirements

Ancess, J. (2000). The reciprocal influence of teacher learning, teaching practice, school restructuring, and student learning outcomes. *Teachers College Record*, *102,* 590-619.

Balfanz, R., Herzog, L., & Mac Iver, D.J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist*, *42,* 223-235.

Balfanz, R. & Legters, N. (2004). Locating the dropout crisis: Which high schools produce the nationôs dropouts? In G. Orfield (Ed.), *Dropouts in America* (pp.57-84). Cambridge, MA: Harvard Education Press.

Belfield, C. & Levin, H. M. Eds. (2007). *The price we pay: Economic and social consequences of inadequate education*. Washington, D.C.: Brookings Institution Press.

Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American schools after a decade*. Santa Monica, CA: RAND.

Berger, A. R., Cole, S., Darwin, M., Duffy, H., Golden, L., Kerem, T., Knudson, J., Rojas, D., Shkolnik, J., Stapleton, J., Stone, C. K., Adelman, N., Cassidy, L., Hall, C., Keating, K., Murray, S., Nielson, N., Park, C. J., Rivera, E., & Yee, K. (2008). *Early college high school initiative evaluation year end report: 2003-2007*. Washington, D.C. and Menlo Park, CA: American Institutes for Research and SRI International. Retrieved July 17, 2009, from: http://www.air.org/publications/pubs_ehd_school_reform.aspx

Bodilly, S. J., Chun, J., Ikemoto, G., & Stockly, S. (2004). *Challenges and potential of a collaborative approach to education reform*. Santa Monica: RAND. Retrieved July 17, 2009, from: http://www.rand.org/pubs/monographs/MG216/

Borman, G.D., Hewes, G.M., Overman, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, *73,* 125-230.

Bridgeland, J. M., DiIulio, J.J. Jr., & Streeter, R. (2008). *Raising the compulsory school attendance age: The case for reform*. Washington, D.C.: Civic Enterprises.

Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good.* Cambridge, MA: Harvard University Press.

Bryk, A. S. & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage.

Cameron, S.V. & Heckman, J.J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics*, *11,* 1-47.

Carnegie Corporation of New York (2004). *Schools for a New Society leads the way*. New York: Author. Retrieved July 17, 2009, from: http://www.carnegie.org/sns/about.html

Carnoy, M., Jacobson, R., Mishel, L., & Rothstein, R. (2005). *The charter school dust-up: Examining the evidence on enrollment and achievement*. Washington, D.C./New York: Economic Policy Institute and Teachers College Press.

Center for Research on Education Outcomes (CREDO) (2009). *Multiple choice: Charter school performance in 16 states. Executive Summary.* Stanford, CA: CREDO, Stanford University.

City University of New York (2009). *College now, high school graduation: Credit and regents requirements*. New York City: Author. Retrieved July 17, 2009, from: http://collegenow.cuny.edu/nextstop/finish_hs/creditreq/

Coburn, C.E. & Talbert, J.E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, *112*, 469-495.

Cohen, D.K. & Spillane, J.P. (1992). Policy and practice: The relations between governance and instruction. *Review of Research in Education*, *18,* 3-49.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd edition. Hillsdale, N.J.: Erlbaum.

Comprehensive School Reform Quality Center (2006). *CSRQ center report on middle and high school comprehensive school reform models*. Washington, D.C.: American Institutes for Research.

Concoran, T. & Silander, M. (2009). Instruction in high schools: The evidence and the challenge. *Future of Children*, *19,* 157-183.

Connell, C., Martin, J., & Moore, L. W. (2002). *The Annenberg Challenge: Lessons and reflections on public school reform*. Philadelphia and Providence: The Annenberg Foundation and the Annenberg Institute for School Reform, Brown University.

Connell, J. P., Legters, N. E., Klem, A., & West, T. C. (2006). *Getting ready, willing and able: Critical steps toward successful implementation of small learning communities in large high schools*. Washington, D.C.: U. S. Department of Education. Retrieved July 17, 2009, from: http://slcp.ed.gov/IssuePapers.html

Cook, T.D. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational Evaluation and Policy Analysis*, *24,* 175-199.

Corrin, W., Somers, M.-A., Kemple, J. J., Nelson, E., & Sepanik, S. (2009). *The enhanced reading opportunities study: Findings from the second year of implementation*. (NCEE 2009-4036) Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Datnow, A. (2000). Power and politics in the adoption of school reform models. *Educational Evaluation and Policy Analysis*, *22,* 357-374.

Deke, J. & Hiamson, J. (2004). *Valuing Student Competencies: Which Ones Predict Postsecondary Educational Attainment and Earnings, and for Whom?* Princeton, NJ: Mathematica. Retrieved July 17, 2009, from: http://www.mathematica-mpr.com/publications/PDFs/valuestudent.pdf

Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, *72,* 433-479.

Dorn, S. (1996). *Creating the dropout: An institutional and social history of school failure*. Westport, CT: Praeger.

Dynarski, M. (2004). Interpreting the evidence from recent federal evaluations of dropout-prevention programs: The state of scientific evidence. In Gary Orfied (Ed.), *Dropouts in America: Confronting the Graduation Rate Crisis* (pp.255-267). Cambridge: Harvard Education Press.

Dynarski, M., Clarke, L., Cobb, B., Finn, J., Rumberger, R., & Smink, J. (2008). *Dropout Prevention: A Practice Guide*. (NCEE 2008-4025). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved July 17, 2009, from: http://ies.ed.gov/ncee/wwc

Dynarski, M. & Gleason, P. (1998). *How can we help? What we have learned from federal dropout-prevention programs*. Princeton, N.J.: Mathematica Policy Research, Inc.

Elmore, R. F. (2004a). Conclusion: The problem of stakes in performance-based accountability systems. In S. H. Furhman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp.274-296). New York: Teachers College Press.

Elmore, R. F. (2004b). *School reform from the inside out*. Cambridge, MA: Harvard Education Press.

Elmore, R. F. & McLaughlin, M. W. (1988). *Steady work: Policy, practice, and the reform of American education*. Santa Monica: The RAND Corporation.

Evan, A., Huberman, M., Means, B., Mitchell, K., Shear/Linda, Shkolnik, J., Smerdon, B., Song, M., Storey, C., & Uekawa, K. (2006). *Evaluation of the Bill and Melinda Gates Foundation's high school grants initiative: 2001-2005 final report*. Washington, D.C. and Menlo Park, CA: American Institutes for Research and SRI International. Retrieved July 17, 2009, from: http://www.air.org/publications/pubs_ehd_school_reform.aspx

Fagan, T. (2008). *Title I funds—Who's gaining, who's losing: School year 2008-09 update*. Washington, D.C.: Center on Education Policy. Retrieved July 17, 2009, from: http://www.cep-dc.org/index.cfm?fuseaction=Page.viewPage&pageId=492&parentID=481

Finn, C. E. Jr., Manno, B. V., & Vanourek, G. (2000). *Charter schools in action: Renewing public education*. Princeton, NJ: Princeton University Press.

Finn, J.D. & Achilles, C.M. (1999). Tennesseeôs class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis*, *21,* 97-110.

Finn, J.D., Gerber, S.B., & Boyd-Zaharias, J. (2005). Small classes in the early grades, academic achievement, and graduating from high school. *Journal of Educational Psychology*, *97,* 214-223.

Fisler, J.L. & Firestone, W.A. (2006). Teacher learning in a school-university partnership: Exploring the role of social trust and teaching efficacy beliefs. *Teachers College Record*, *108*, 1155-1185.

Fleischman, S. & Heppen, J. (2009). Improving low-performing high schools: Searching for evidence of promise. *Future of Children*, *19,* 105-133.

Foley, E. M., Klinge, A., & Reisner, E. R. (2008). *Evaluation of New Century High Schools: Profile of an initiative to create and sustain small, successful high schools*. New York:

Policy Studies Associates. Retrieved July 17, 2009, from: http://www.policystudies.com/studies/school/New%20Century%20year%201.html

Friedlaender, D. & Darling-Hammon, L. (2007). *High schools for equity: Policy supports for student learning in communities of color*. Stanford, CA: School Redesign Network at Stanford University. Retrieved April 12, 2009, from: http://www.srnleads.org/resources/publications/hsfe.html

Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century*. New York: Farrar, Straus, & Giroux.

Furhman, S. H. & Elmore, R. F. (Eds.) (2004). *Redesigning accountability systems for education*. New York: Teachers College Press.

Gándara, P. & Bial, D. (2001). *Paving the way to postsecondary education: K-12 intervention programs for underrepresented youth*. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics. Retrieved July 17, 2009, from: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2001205

General Education Development Testing Service (2008). *2007 GED Testing Program Statistical Report*. Washington, D.C.: Author.

Glennan, T. K. Jr., Bodilly, S. J., Galegher, J. R., & Kerr, K. A. (2004). *Expanding the reach of education reforms: Perspectives from leaders in the scale-up of educational interventions*. Santa Monica, CA: RAND.

Greene, J. P. & Winters, M. A. (2005). *Public school graduation and college-readiness rates: 1991-2002*. New York: Manhattan Institute.

Grigg, W., Donahue, P. L., & Dion, G. (2007). *The nation's report card: 12th-grade reading and mathematics 2005*. (NCES 2007-468). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics. Retrieved July 17, 2009, from: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007468

Grubb, W. N. (2009). *The money myth: School resources, outcomes, and equity*. New York: Russell Sage.

Guskey, T.R. (1989). Attitude and perceptual change in teachers. *International Journal of Educational Research*, *13,* 439-453.

Hanushek, E. A. (1996). Outcomes, costs, and incentives in schools. In E. A. Hanushek & D. W. Jorgenson (Eds.), *Improving America's schools: The role of Incentives* (pp.29-52). Washington, D.C.: National Academy Press.

Hanushek, E. A. & Jorgenson, D. W. (Eds.) (1996). *Improving America's schools: The role of incentives*. Washington, D.C.: National Academy Press.

Harr, J. J., Parrish, T., Socias, M., & Gubbins, P. (2007). *Evaluation study of California's High Priority Schools Grant Program: Final report*. Palo Alto, CA: American Institutes for Research. Retrieved July 17, 2009, from: http://www.air.org/publications/pubs_ehd_school_reform.aspx

Heckman, J. J. & LaFontaine, P. A. (2008). *The American high school graduation rate: Trends and levels*. NBER Working Paper 13670. Cambridge, MA: National Bureau of Economic Research. Retrieved July 17, 2009, from http://www.nber.org/papers/w13670

Heckman, J.J. & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91,* 145-149.

Henig, J. R. (2008). *Spin cycle: How research is used in policy debate:The case of charter schools*. New York: Russell Sage Foundation.

Herlihy, C. & Quint, J. (2007). *Emerging evidence on improving high school student achievement and graduation rates: The effects of four popular improvement programs*. Washington, D.C.: National High School Center, American Institutes for Research. Retrieved July 17, 2009, from http://www.betterhighschools.org/pubs/

Herman, J. L. & Baker, E. L. (2003). *The Los Angeles Annenberg Metropolitan Project: Evaluation findings*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing. Retrieved July 17, 2009, from: http://www.cse.ucla.edu/products/summary.asp?report=591

Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning around chronically low-performing schools*. (NCEE 2008-4020). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved July 17, 2009, from: http://ies.ed.gov/ncee/wwc

Hoffman, L. (2008). *Numbers and types of public elementary and secondary education agencies from the Common Core of Data: School year 2006-07*. (NCES 2009-304REV). Washington, D.C.: U.S. Department of Labor, National Center for Education Statistics. Retrieved July 17, 2009, from: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2009303

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices*. (NCEE 2008-4027). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved July 17, 2009, from: http://ies.ed.gov/ncee/wwc

Kelly, D. M. (1993). *Last chance high: How girls and boys drop in and out of alternative schools*. New Haven: Yale University Press.

Krueger, C. (2006). *Dual enrollment: Policy issues confronting state policymakers*. Policy Brief. Denver: Education Commission of the States.

Laird, J., Cataldi, E. F., KewalRamani, A., & Chapman, C. (2008). *Dropout and completion rates in the United States: 2006*. (NCES 2008-043). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics. Retrieved July 17, 2009, from: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2008053

Lee, V.E. & Bryk, A.S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education*, *62,* 172-192.

Lee, V.E. & Smith, J.B. (1995). Effects of high school restructuring and size on gains in achievement and engagement for early secondary school students. *Sociology of Education*, *68,* 241-279.

Lee, V.E. & Smith, J.B. (1997). High school size: Which works best for whom? *Educational Evaluation and Policy Analysis*, *19,* 205-227.

Lee, V. E. & Smith, J. B. (2001). *Restructuring high schools for equity and excellence*. New York: Teachers College Press.

Lehr, C. A., Lanners, E. J., & Lange, C. M. (2003). *Alternative schools: Policy and legislation across the United States*. Minneapolis, MN: University of Minnesota, Institute on Community Integration. Retrieved July 17, 2009, from: http://ici.umn.edu/alternativeschools/publications/default.html

Levin, H. (1988). Cost-effectiveness and educational policy. *Educational Evaluation and Policy Analysis*, *10,* 51-69.

Lippman, L., Atienza, A., Rivers, A., & Keith, J. (2008). *A developmental perspective on college and workplace readiness*. Washington, D.C.: Child Trends. Retrieved July 17, 2009, from: http://www.childtrends.org/_listES.cfm?LID=C33B0F7B-0ACC-41C4-A7E8C8B675B22EF6#Education

Little, J.W. (2003). Inside teacher community: Representations of classroom practice. *Teachers College Record*, *105,* 913-945.

Louis, K.S., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachersô interpretations of a new era. *Educational Evaluation and Policy Analysis*, *27,* 177-204.

Louis, K.S. & Marks, H.M. (1998). Does professional community affect the classroom? Teachersô work and student experiences in restructuring schools. *American Journal of Education*, *106,* 532-575.

Louis, K.S., Marks, H.M., & Kruse, S. (1996). Teachersô professional community in restructuring schools. *American Educational Research Journal*, *33,* 757-798.

Malen, B. & Rice, J.K. (2004). A framework for assessing the impact of education reforms on school capacity: Insights from studies of high-stakes accountability initiatives. *Educational Policy*, *18,* 631-660.

Mathis, W. J. (2009). *NCLB's ultimate restructuring alternatives: Do they improve the quality of education*. Boulder, CO and Tempe, AZ: Education and the Public Interest Center & Education Policy Research Unit. Retrieved July 17, 2009, from http://epicpolicy.org/publication/nclb-ultimate-restructuring

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND.

McDonnell, L. M. (2004). *Politics, persuasion, and educational testing*. Cambridge, MA: Harvard University Press.

McDonnell, L.M. & Elmore, R.F. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis*, *9,* 133-152.

McLaughlin, M.W. (1990). The Rand Change Agent Study revisited: Macro perspectives and micro realities. *Educational Researcher*, *19,* 11-16.

McLaughlin, M. W. & Talbert, J. E. (2001). *Professional communities and the work of high school teaching*. Chicago: University of Chicago Press.

McLaughlin, M.W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, *9,* 171-178.

Meeder, H. (2008). *The Perkins Act of 2006: Connecting career and technical education with the college and career readiness agenda*. Washington, D.C.: Achieve, Inc.

Miao, J. & Haney, W. (2004). High school graduation rates: Alternative methods and implications. *Education Policy Analysis Archives*, *12,* Retrieved July 17, 2009, from: http://epaa.asu.edu/epaa/v12n55/

Minnici, A. & Hill, D. D. (2007). *Educational architects: Do state education agencies have the tools necessary to implement NCLB?* Washington, D.C.: Center on Education Policy. Retrieved July 17, 2009, from: http://www.cep-dc.org/index.cfm?fuseaction=Page.viewPage&pageId=494&parentID=481

National Research Council, Committee on Appropriate Test Use (1999). *High stakes: Testing for tracking, promotion, and graduation*. Jay P. Heubert and Robert M. Hauser (Eds.). Washington, D.C.: National Academies Press.

National Research Council, Committee on Increasing High School Students' Engagement and Motivation to Learn (2004). *Engaging Schools: Fostering High School Students' Motivation to Learn*. Washington, D.C.: The National Academies Press.

National Research Council, Panel on High-Risk Youth (1993). *Losing generations: Adolescents in high-risk settings*. Washington, D.C.: National Academies Press.

Newmann, F.M. (1993). Beyond common sense in educational restructuring. *Educational Researcher*, *22,* 4-13, 22.

Newmann, F.M., King, M.B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review*, *67,* 41-74.

Newmann, F.M., King, M.B., & Youngs, P. (2000). Professional development that addresses school capacity: Lessons from urban elementary schools. *American Journal of Education*, *108,* 259-299.

Newmann, F.M., Smith, B., Allensworth, E., & Bryk, A.S. (2001). Instructional program coherence: What it is and why it should guide school improvement policy. *Educational Evaluation and Policy Analysis*, *23,* 297-321.

Olson, L. (June 12, 2007). What Does ñReadyò Mean? *Education Week.* Retrieved July 17, 2009, from: http://www.edweek.org/ew/articles/2007-06/12/40overview.h26.html

Orfield, G. & Lee, C. (2006). *Racial transformation and the changing nature of segregation*. Cambridge, MA: The Civil Rights Project, Harvard University.

Organization for Economic Co-operation and Development (2001). *Knowledge and skills for life: First results from the OECD programme for international student assessment (PISA) 2000*. Paris: OECD.

Organization for Economic Co-operation and Development (2008). *Education at a Glance 2008*. Paris: OECD.

Peter D. Hart Research Associates/Public Opinion Strategies (2005). *Rising to the challenge: Are high school graduates prepared for college and work?* Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.achieve.org/RisingtotheChallenge

Planty, M., Provasnik, S., & Daniel, B. (2007). *High school coursetaking: Findings from the Condition of Education 2007*. (NCES 2007-065) Washington, D.C.: National Center for Education Statistics.

Porter, A.C. (1989). External standards and good teaching: The pros and cons of telling teachers what to do. *Educational Evaluation and Policy Analysis*, *11,* 343-356.

Porter, K. E. & Snipes, J. C. (2006). *The challenge of supporting change: Elementary student achievement and the Bary Area School Reform Collaborative's focal strategy*. New York: MDRC. Retrieved July 17, 2009, from: http://www.mdrc.org/publications/439/execsum.html

Purkey, S.C. & Smith, M.S. (1983). Effective schools: A review. *The Elementary School Journal*, *83,* 426-452.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. 2nd edition. Thousand Oaks, CA: Sage Publications.

Raudenbush, S.W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, *29,* 5-29.

Reardon, S. F., Atteberry, A., Arshan, N., & Kurlaender, M. (2009). *Effects of California High School Exit Exam on student persistence, achievement, and graduation*. Stanford, CA: Institute for Research on Education Policy & Practice, Stanford University. Retrieved July 17, 2009, from: http://www.stanford.edu/group/irepp/cgi-bin/joomla/index.php

Reyes, P., Alexander, C., Fuller, E., & Phillips, J. C. (2007). *Houston schools for a new society*. Austin: Study of High School Restructuring, University of Texas. Retrieved July 17, 2009, from: http://www.edb.utexas.edu/hsns/publication.html

Roderick, M., Easton, J. Q., & Sebring, P. B. (2009). *The Consortium on Chicago School Research: A new model for the role of research in supporting urban school reform*. Chicago: University of Chicago Urban Education Institute, The Consortium on Chicago School Research. Retrieved July 17, 2009, from: http://ccsr.uchicago.edu/content/publications.php?pub_id=131

Rogers, J., Fanelli, S., Medina, D., Zhu, Q., Freelon, R., Bertrand, M., & Del Razo, J. (2009). *California educational opportunity report: Listening to public school parents*. Los Angeles: UCLA/IDEA and UC/ACCORD.

Rotermund, S. (2007). *Alternative education enrollment and dropouts in California high schools*. Statistical Brief 6. Santa Barbara: California Dropout Research Project, University of California, Santa Barbara. Retrieved July 17, 2009, from http://lmri.ucsb.edu/dropouts/pubs.htm

Rothstein, R., Jacobsen, R., & Wilder, T. (2008). *Grading education: Getting accountability right*. Washington, D.C. and New York, NY: Economic Policy Institute and Teachers College Press.

Rumberger, R. W. (2008). Can NCLB improve high school graduation rates? In G. L. Sunderman (Ed.), *Holding NCLB accountable: Achieving accountability, equity, and school reform* (pp.209-222). Thousand Oaks, CA: Corwin Press.

Rumberger, R. W. (forthcoming). *Why students drop out of school and what can be done about it (working title)*. Cambridge, MA: Harvard University Press.

Rumberger, R. W. & Connell, J. (2007). *Strengthening school district capacity as a strategy to raise student achievement in California*. Mountain View, CA: EdSource.

Rumberger, R. W. & Lim, S. A. (2008). *Why students drop out of school: A review of 25 years of research*. Santa Barbara, CA: California Dropout Research Project. Retrieved July 17, 2009, from: http://lmri.ucsb.edu/dropouts/publications

Rumberger, R. W. & Palardy, G. J. (2004). Multilevel models for school effectiveness research. In D. Kaplan (Ed.), *Handbook of Quantitative Methodology for the Social Sciences* (pp.235-258). Thousand Oaks, CA: Sage Publications.

Rumberger, R.W. & Palardy, G.J. (2005a). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, *107,* 1999-2045.

Rumberger, R.W. & Palardy, G.J. (2005b). Test scores, dropout rates, and transfer rates as alternative indicators of high school performance. *American Educational Research Journal*, *41,* 3-42.

Rumberger, R. W. & Rotermund, S. (2008). *What happened to dropouts from the high school class of 2004?* Statistical Brief 10. Santa Barbara: California Dropout Research Project, University of California, Santa Barbara. Retrieved July 17, 2009, from http://lmri.ucsb.edu/dropouts/pubs.htm

Sarason, S. B. (1996). *Revisiting "the culture of the school and the problem of change."* New York: Teachers College Press.

Schneider, B., Carnoy, M., Kilpatrick, J.,Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs*. Report from the Governing Board of the American Educational Research Association Grants Program. Washington, D.C.: American Educational Research Association.

Shavelson, R.J., McDonnell, L., Oakes, J., & Carey, N. (1987). *Indicator Systems for Monitoring Mathematics and Science Education*. Santa Monica: RAND.

Shettle, C., Roey, S., Mordica, J., Perkins, R., Nord , C., Teodorovic, J., Lyons, M., Averett, C., Kastberg, D., & Brown, J. (2007). *America's high school graduates:Results from the 2005 NAEP high school transcript study*. (NCES 2007-467). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Slavin, R.E. (2008). What works? Issues in synthesizing educational program evaluations. *Educational Researcher*, *37,* 5-14.

Slavin, R.E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly*, *43,* 290-322.

Slavin, R.E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, *Retrieved July 17, 2009,* from: http://rer.sagepub.com.proxy.library.ucsb.edu:2048/cgi/rapidpdf/0034654308330968v1

Smith, M. S. & OôDay, J. (1991). Systemic school reform. In S. H. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp.233-267). Philadelphia: Falmer Press.

Snyder, T. D., Dillow, S. A., & Hoffman, C. M. (2009). *Digest of Education Statistics, 2008*. (NCES 2008-022) U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office. Retrieved July 17, 2009, from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2009020

Spillane, J.P. (1996). School districts matter: Local educational authorities and state instructional policy. *Educational Policy*, *10*, 63-87.

Stecher, B. M., Epstein, S., Hamilton, L. S., Marsh, J. A., Robyn, A., McCombs, J. S., Russell, J., & Naftel, S. (2008). *Pain and gain: Implementing No Child Left Behind in three states, 2004-2006*. Santa Monica: RAND.

Sunderman, G.L. & Kim, J.S. (2007). The expansion of federal power and the politics of implementing the No Child Left Behind Act. *Teachers College Record*, *109,* 1057-1085.

Sunderman, G. L. & Orfied, G. (2008). Massive responsibilities and limited resources: The state response to NCLB. In G. L. Sunderman (Ed.), *Holding NCLB accountable: Achieving accountability, equity, and school reform* (pp.121-136.). Thousand Oaks, CA: Corwin Press.

Supovitz, J. A. (2006). *The case for district-based reform: Leading, building, and sustaining school improvement*. Cambridge, MA: Harvard Education Press.

Texas Education Agency (2008). *Grade Level Retention in Texas Public Schools, 2006-07*. Austin: Texas Education Agency. Retrieved July 17, 2009, from: http://ritter.tea.state.tx.us/research/pdfs/retention_2006-07.pdf

Tyack, D. B. (1974). *The one best system*. Cambridge, MA: Harvard University Press.

Tyler, J.H. (2003). Economic benefits of the GED: Lessons from recent research. *Review of Educational Research*, *73,* 369-398.

U.S. Department of Education (2002). *Comprehensive School Reform (CSR) Program Guidance*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.ed.gov/programs/compreform/legislation.html

U.S. Department of Education (2008). *A uniform, comparable graduation rate*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.ed.gov/policy/elsec/reg/proposal/uniform-grad-rate.html

U.S. Department of Education (2009). *American Recovery and Reinvestment Act of 2009: Using ARRA funds to drive school reform and improvement*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.ed.gov/policy/gen/leg/recovery/index.html

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2008). *WWC topic report: Dropout prevention*. Retrieved July 17, 2009, from: http://ies.ed.gov/ncee/wwc/reports/topic.aspx?tid=06

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2009). *WWC intervention report: SuccessMaker*. Retrieved July 17, 2009, from: http://ies.ed.gov/ncee/wwc/reports/topic.aspx?tid=06

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service. (2008a). *Evaluation of the Comprehensive School Reform Program implementation and outcomes: Third-year report*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.ed.gov/about/offices/list/opepd/ppss/reports.html

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service. (2008b). *Implementation study of smaller learning communities, final report*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.ed.gov/about/offices/list/opepd/ppss/reports.html

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service. (2009). *Title I implementation—Update on recent evaluation findings*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.ed.gov/about/offices/list/opepd/ppss/reports.html

U.S. Department of Education, Office of the Under Secretary, Policy and Program Studies Service. (2004). *National assessment of vocational education: Final report to Congress*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.ed.gov/about/offices/list/opepd/ppss/reports.html

U.S. General Accountability Office (2005). *No Child Left Behind Act: Education could do more to help states better define graduation rates and improve knowledge about intervention strategies*. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.gao.gov/products/GAO-05-879.

U.S. President. (2009). The White House: Briefing Room. Washington, D.C.: Author. Retrieved July 17, 2009, from: http://www.whitehouse.gov/the_press_office/Fact-Sheet-Expanding-the-Promise-of-Education-in-America/

Wagner, T. (2008). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need—and what we can do about it*. New York: Basic Books.

Warren, J.R. (2005). State-level high school completion rates: Concepts, measures, and trends. *Education Policy Analysis Archives*, *13,* Retrieved July 17, 2009, from: http://epaa.asu.edu/epaa/v13n51/

Warren, J.R., Jenkins, K.N., & Kulick, R.B. (2006). High school exit examinations and state-level completion and GED rates, 1975-2002. *Educational Evaluation and Policy Analysis*, *28,* 131-152.

Wehlage, G., Smith, G., & Lipman, P. (1992). Restructuring urban schools: The New Futures experience. *American Educational Research Journal*, *29,* 51-93.

White, J.A. & Wehlage, G. (1995). Community collaboration: If it is such a good idea, why is it so hard to do? *Educational Evaluation and Policy Analysis*, *17,* 23-38.

Woodworth, K. R., David, J. L., Guha, R., Wang, H., & Lopez-Torkos, A. (2008). *San Francisco Bay Area KIPP schools: A study of early implementation and achievement. Final report.* Menlo Park, CA: SRI International. Retrieved July 17, 2009, from: http://policyweb.sri.com/cep/projects/displayProject.jsp?Nick=kipp

Zimmer, R., Gill, B., Booker, K., Lavertu, S., Sass, T. R., & Witte, J. (2009). *Charter schools in eight states: Effects on achievement, attainment, integration, and competition*. Santa Monica, CA: Rand.

Table 1. Relative Importance of Eight Goals of Public Education Rated by Three Groups of Respondents

	General public	School board members	State legislators
Basic academic skills	19%	22%	24%
Critical thinking	15%	18%	18%
The arts and literature	8%	9%	9%
Preparation for skilled work	11%	11%	11%
Social skills and work ethic	14%	12%	11%
Citizenship	10%	11%	12%
Physical health	12%	9%	9%
Emotional health	11%	8%	7%
TOTAL	**100**	**100**	**100**

SOURCE: Rothstein, Jacobsen, & Wilder (2008), Table 1.

Table 2. Federal High School Improvement Programs

Name	Description	Funding
Vocational Education—Basic Grants to States http://www.ed.gov/programs/ctesbg/index.html	The purpose of the program is to develop more fully the academic and career and technical skills of secondary and postsecondary students who elect to enroll in career and technical programs. The program provides states with support for state leadership activities, administration of the state plan for vocational and technical education, and subgrants to eligible recipients to improve vocational and technical education programs, including the use of technology and professional development.	FY2008 $1.1 B FY2007 $1.2 B FY2006 $1.2 B FY2005 $1.2 B FY2004 $1.2 B FY2003 $1.2 B FY2002 $1.2 B FY2001 $1.1 B FY2000 $1.0 B
Comprehensive School Reform Program http://www.ed.gov/programs/compreform/index.html	Provides formula grants to State Educational Agencies (SEAs) to award competitive grants to Local Educational Agencies (LEAs) on behalf of schools and a grant for CSR clearinghouse. The grants are designed to raise student achievement by assisting public schools across the country to implement effective, comprehensive school reforms founded upon scientifically based research and effective practices.	FY2008 $1.6 M FY2007 $1.5 M FY2006 $1.5 M FY2005 $205 M FY2004 $308 M FY2003 $308 M FY2002 $310 M FY2001 $260 M FY2000 $220 M
Comprehensive School Reform Quality Initiatives http://www.ed.gov/programs/qualinits/index.html	The purpose of the CSR Quality Initiatives program is to provide discretionary grants to support activities that will enhance the state-administered CSR program and to enable schools that have been identified for improvement, corrective action, or restructuring under Title I, Part A of ESEA to meet their stateôs definition of adequate yearly progress.	FY2007 $0.8 M FY2006 $6.5 M FY2005 $7.1 M FY2004 $7.1 M
Comprehensive Centers Program http://www.ed.gov/programs/newccp/index.html	This program supports 21 comprehensive centers to help increase state capacity to help districts and schools meet their student achievement goals; 16 regional centers that provide services primarily to state education agencies (SEAs) to enable them to assist school districts, regional education agencies, and schools; and five content centers for accountability, instruction, teacher quality, innovation and improvement, and high schools that will supply much of the research-based information and products in the specific area that regional centers will use when working with states.	FY2008 $57 M FY2007 $56 M FY2006 $56 M FY2005 $57 M

School Improvement Fund http://www.ed.gov/programs/sif/index.html	Provides formula grants to State Educational Agencies (SEAs) to assist schools identified for improvement, corrective action, and restructuring by building state capacity to provide leadership in implementing effective school improvement strategies for Local Educational Agencies (LEAs) and schools. The Fund also provides resources to LEAs to support school improvement activities, including the development and implementation of effective school improvement plans.	RA2009 $3.0 B* FY2008 $491 M FY2007 $125 M
School Dropout Prevention Program http://www.ed.gov/programs/dropout/index.html	Supports effective, sustainable, and coordinated dropout prevention and reentry programs in high schools with annual dropout rates that exceed their state average annual dropout rate. Middle schools that have students who continue on to these high schools also are supported.	FY2006 $4.9 M FY2005 $4.9 M FY2004 $5.0 M
Smaller Learning Communities Program http://www.ed.gov/programs/slcp/index.html	Awards discretionary grants to local educational agencies (LEAs) to support the implementation of SLCs and activities to improve student academic achievement in large public high schools with enrollments of 1,000 or more students. SLCs include structures such as freshman academies, multi-grade academies organized around career interests or other themes, ñhousesò in which small groups of students remain together throughout high school, and autonomous schools-within-a-school, as well as personalization strategies, such as student advisories, family advocate systems, and mentoring programs.	FY2008 $80 M FY2007 $94 M FY2006 $94 M FY2005 $94 M FY2004 $174 M FY2003 $161 M FY2002 $142 M FY2001 $125 M FY2000 $45 M
Striving Readers Program http://www.ed.gov/programs/strivingreaders/index.html	Aims to raise middle and high school studentsô literacy levels in Title I-eligible schools with significant numbers of students reading below grade levels and to build a strong, scientific research base for strategies that improve adolescent literacy instruction by funding projects that provide a supplemental literacy intervention targeted to low-performing students and a school-wide literacy program for improving literacy of all students	FY2008 $35 M FY2007 $32 M FY2006 $39 M FY2005 $25 M

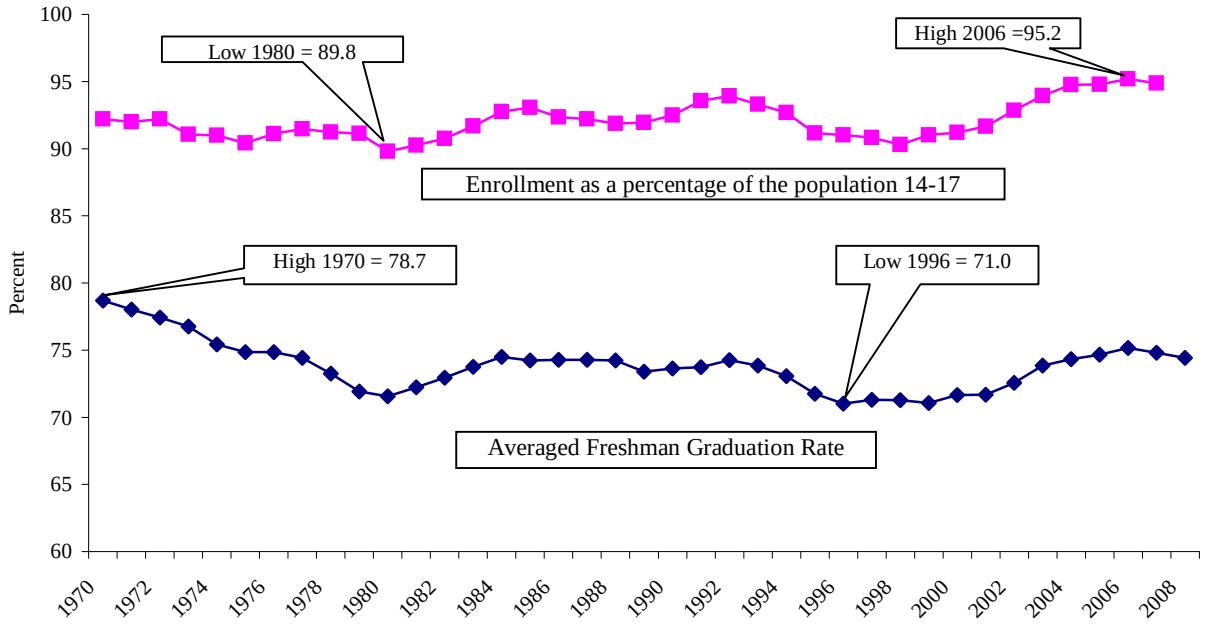
*Funding from the American Recovery and Reinvestment Act of 2009 (Recovery Act) that was signed into law by President Obama on February 17, 2009. See: http://www.ed.gov/policy/gen/leg/recovery/index.html

Table 3. Features of Policy Instruments

Instrument	Mechanism	Policy Problem	Assumptions and Expected Effects
Mandate	Rules	Undesirable behavior or goods being produced; Lack of uniform standards	Have capacity to comply; Most will do so, although some shirking likely
Inducement	Money for procurement	Valued goods or services not being produced with desired frequency or quality	Capacity exists or easily acquired; Money will elicit performance, although variability likely
Capacity-building	Money for investment	Lack of longer-term investment in needed skills and valued goods	Capacity does not exist; Knowledge and skills required to produce future value
System-changing	Authority	Existing institutional arrangements not producing desired outcomes	New entrants will produce desired results and will motivate established institutions to improve; New entrants may generate other problems.
Hortatory	Information Beliefs and values	Information for remedying a problem or making a choice is incomplete; Some values held strongly, so beyond the control of just incentives or rules.	Educators will respond positively because they share values and goals of policymakers

SOURCE: Adapted from McDonnell (2004), Table 2.1.

Figure 1. Enrollment and Graduation Rates, 1970-2008



SOURCE: (Rumberger, forthcoming, Table 1.6).