NATIONAL CENTER FOR EDUCATION STATISTICS

Statistical Analysis Report December 2000

MONITORING SCHOOL QUALITY: An Indicators Report

U.S. Department of Education
Office of Educational Research and Improvement NCES 2001–030

NATIONAL CENTER FOR EDUCATION STATISTICS

Statistical Analysis Report December 2000

MONITORING SCHOOL QUALITY: An Indicators Report

Daniel P. Mayer
John E. Mullens
Mary T. Moore
Mathematica Policy Research, Inc.

John Ralph, Project Officer
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement NCES 2001–030

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

> National Center for Education Statistics
> Office of Educational Research and Improvement
> U.S. Department of Education
> 1990 K Street, NW
> Washington, DC 20006-5574

December 2000

The NCES World Wide Web Home Page is: ***http://nces.ed.gov***
The NCES World Wide Web Electronic Catalog is: ***http://nces.ed.gov/pubsearch/index.asp***

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. Monitoring School Quality: An Indicators Report, NCES 2001–030 by Daniel P. Mayer, John E. Mullens, and Mary T. Moore. John Ralph, Project Officer. Washington, DC: 2000.

For ordering information on this report, write:

> U.S. Department of Education
> ED Pubs
> 8242-B Sandy Court
> Jessup, MD 20794–1398

or call toll free 1–887–4ED–Pubs.

Content Contact:
John Ralph
(202) 502–7441

ACKNOWLEDGEMENTS

This work could not have been accomplished without the guidance, consultation, and support of many people. John Ralph, the project officer from the National Center for Education Statistics (NCES), U.S. Department of Education asked thoughtful, penetrating questions and provided valuable advice and guidance from the study's conception to its fruition. Also at NCES, Marty Orland, while Associate Commissioner, played a key role in conceptualizing the project to which Val Plisko, now Associate Commissioner for Early Childhood, International and Crosscutting Studies, added her insightful perspective.

Several people contributed to the conception and the final product. NCES reviewers Marilyn McMillen, Shelley Burns, John Wirt and Ellen Bradburn provided detailed, discerning, and important comments. External advisors and reviewers provided us with helpful comments that helped sharpen the report's focus. For that we wish to thank Paul Barton, Educational Testing Service; Joyce Benjamin, Oregon Department of Education; Lorraine McDonnell, University of California, Santa Barbara; Jack Jennings, Center on Education Policy; Richard Murnane, Harvard University Graduate School of Education; and Meredith Phillips, UCLA School of Public Policy and Social Research.

At Mathematica Policy Research, David Myers provided critical suggestions concerning the study's direction and Paul Decker, author of the first indicator study in this series, contributed his unique perspective. Julia Kim was an early source of much background research that laid the groundwork for the completed report. Heather Hesketh and Emily Pas helped collect data and produce the tables for the report, Daryl Hall and Anne Kelleher helped edit it, and Felita Buckner and Alfreda Holmes helped produce the report. To each person, we extend our gratitude and appreciation.

CONTENTS

TABLES

FIGURES

EXECUTIVE SUMMARY

This report explores why some schools may be better than others at helping students learn. It responds to a recommendation from the congressionally mandated Special Study Panel on Education Indicators for the National Center for Education Statistics (NCES) for reports that identify and discuss indicators of the health of the nation's educational system (U.S. Department of Education 1991). This report is designed for policymakers, researchers, and others interested in assessing the strength of our schools. While it is relevant for those interested in standards or accountability, it is not about test scores and is not a guide for education reform movements.

More specifically, the report's primary goals are to

1. Review the literature on school quality to help policymakers and researchers understand what is known about the characteristics of schools that are most likely related to student learning,
2. Identify where national indicator data are currently available and reliable, and
3. Assess the current status of our schools by examining and critiquing these national indicator data.

SCHOOL CHARACTERISTICS RELATED TO STUDENT LEARNING

The research described in this report indicates that school quality affects student learning through the training and talent of the teaching force, what goes on in the classrooms, and the overall culture and atmosphere of the school. Within these three areas, this report identifies 13 indicators of school quality that recent research suggests are related to student learning and reviews the national data showing the current status of our schools. These indicators, and the quality of the data describing each, are summarized in Figure ES.1. The figure illustrates that these school quality factors can affect student learning *both directly and indirectly.* For example, school context characteristics like school leadership can have an impact on teachers and what they are able to accomplish in the classroom, and this in turn may influence student learning. In addition, various teacher-level attributes can affect the quality of the classroom and in turn student learning. Traits at each of these levels can also directly affect student learning.

Teachers

Substantial research suggests that school quality is enhanced when teachers have high academic skills, teach in the field in which they are trained, have more than a few years of experience, and participate in high-quality induction and professional development programs. Students learn more from teachers with strong academic skills and classroom teaching experience than they do from teachers with weak academic skills and less experience. Teachers are less effective in terms of student outcomes when they teach cours-

Figure ES.1—School quality indicators and their relationship to student learning

es they were not trained to teach. Teachers are thought to be more effective when they have participated in quality professional development activities, but there is no statistical evidence to evaluate this relationship.

Classrooms

To understand the effectiveness of classrooms, research suggests that it is necessary to understand the content of the curriculum; the pedagogy, materials, and equipment used. Students appear to benefit when course content is focused and has a high level of intellectual rigor and cognitive challenge. Younger students, especially disadvantaged and minority students, appear to learn better in smaller classes. Nationally representative data on the process of schooling, now becoming available for the first time, will further our understanding of the role of these factors in determining school quality.

School Context

How schools approach educational leadership and school goals, develop a professional community, and establish a climate that minimizes discipline problems and encourages academic excellence clearly affects school quality and student learning. For three reasons, however, the effect of school-level characteristics is more difficult to ascertain than the effect of teachers and classrooms. First, even though they are integral to a school, these characteristics are difficult to define and measure. Second, their effect on student learning is likely to be exerted indirectly through teachers and classrooms, compounding the measurement problem. And last, with some exceptions, reliable school-representative information about these indicators of quality is minimal. These difficulties should not overshadow the importance of collecting such data to learn more about how these characteristics operate and affect student learning through teachers and classrooms. The preponderance of national, regional, and local efforts to develop quality schools heightens the benefits that would be derived from additional refined and reliable school-representative measures of school characteristics.

AVAILABILITY AND QUALITY OF INDICATOR DATA

The quality of existing data on these three types of indicators varies (Table ES.1). Where the dimension being measured is straightforward, or if it has been measured for an extended period of time, the data are high quality. Where there is little information about a particular important facet of an indicator, the data quality is moderated in some aspect. And where the indicator is more complex than the data, the quality is poor. For a few indicators, concrete statistical evidence of an association with learning is thin, even though experts agree that these indicators should enhance student learning.

The indicators of teacher assignment, teacher experience, and class size each represent straightforward concepts and are easy to measure, and the data on these indicators are high quality. In addition, data on teacher experience and class size have been collected over several decades, further ensuring their quality. Data on teacher academic skills are also high quality, albeit less straightforward. While the academic skills of teachers are only one aspect of teaching ability, standardized tests that measure the academic skills of future teachers are quite advanced and have consistently shown links to student learning.

Data on indicators of professional development, course content, technology, discipline, and academic organization are moderate in quality. National data-collection efforts pertaining to these indicators are relatively new, and these dimensions of schools are more complex than the data currently collected. Consequently, data on professional development are limited and provide little insight into important principles of successful professional development programs. National data on indicators of course content and academic environment are based primarily on course titles and are consequently too vague to be high quality. Current data on technology primarily measure the availability of hardware and access to the Internet and provide too little information on the instructional role of technology in the classroom. Nationally representative data on school discipline incidents and on school discipline policies are well defined, but administrators may underreport their discipline problems. In addition, there are limited data documenting a link to student learning, the implementation of discipline policies, and their perceived fairness.

Only poor-quality data are available on teachers' pedagogy, school leadership, school goals, and professional community. These indicators are complex and therefore more difficult to measure, and historically they have not been prominent in national data-collection efforts. It is difficult to isolate and measure critical elements of pedagogy because the teaching process consists of a complex set of interactions between students,

Table ES.1—Quality of national school quality indicator data

High Quality	Moderate Quality	Poor Quality
Teacher Assignment	Professional Development	Pedagogy
Teacher Experience	Technology	Goals
Teacher Academic Skills	Course Content	School Leadership
Class Size	Discipline	Professional Community
	Academic Environment	

the teacher, and the curriculum. Measuring human actions, incentives, and opinions to estimate the effects of school-level attributes such as leadership, goals, and professional community is an equally complex task.

As a group, the teacher-focused measures of school quality are less complex and have been collected for some time. School-level attributes of quality are nearly the opposite. We have more reliable information on indicators with high-quality data, while indicators with lower-quality data provide an incentive and direction for improved national data collection. Nine indicators have high- or moderate-quality data and describe the current status of school quality.

CURRENT STATUS OF SCHOOL QUALITY

The Academic Skills of Teachers

Students learn more from teachers with strong academic skills (Ballou 1996; Ehrenberg and Brewer 1994; 1995; Ferguson 1991; Ferguson and Ladd 1996; Mosteller and Moynihan 1972), but graduates whose college entrance examination scores are in the top quartile are *half* as likely as those in the bottom quartile to prepare to teach (9 versus 18 percent) (Henke, Chen, and Geis 2000). Teachers in the top quartile are more than *twice* as likely than teachers in the bottom quartile to teach in private schools (26 versus 10 percent) and are less than *one-third* as likely as teachers in the bottom quartile to teach in high-poverty schools (10 versus 31 percent). Furthermore, graduates in the top quartile who teach are *twice* as likely as those in the bottom quartile to leave the profession within less than four years (32 versus 16 percent) (Henke, Chen, and Geis 2000).

Teaching Assignment

Middle and high school students learn more from teachers who hold a bachelor's or master's degree in the subject they are teaching (Darling-Hammond 2000; Goldhaber and Brewer 1997; Monk and King 1994), but out-of-field teaching occurs with regularity (Bobbitt and McMillen 1994; Henke et al. 1997; Ingersoll 1999; Lewis et al. 1999).

Teacher Experience

Studies suggest that students learn more from experienced teachers than they do from less experienced teachers (Darling-Hammond 2000; Murnane and Phillips 1981; Rivkin, Hanushek, and Kain 1998). As of 1998, the highest-poverty schools and schools with the highest concentrations of minority students had nearly *double* the proportion of inexperienced teachers (those with three or fewer years of experience) than schools with the lowest poverty (20 versus 11 percent) and lowest concentration of minority students (21 versus 10 percent).

Professional Development

Experts agree that high-quality professional development should enhance student learning (Choy and Ross 1998; Mullens et al. 1996; U.S. Department of Education 1999a), but data permitting an analysis of the relationship are not yet available. In 1998, 99 percent of the nation's public school teachers had participated in some type of professional development program within the past 12 months (U.S. Department of

Education 1999). However, most teachers participated in these activities from only one to eight hours, or for no more than one day. Teachers with three or fewer years of experience were more likely (Lewis et al. 1999) to have reported participating in an induction program in 1998–99 than in 1993–94 (65 versus 59 percent).

Course Content

Research shows that as students take higher-level academic courses they learn more (Raizen and Jones 1985; Sebring 1987). From 1982 to 1998, there was an increase in the percentage of students enrolling in higher-level mathematics and science courses (Wirt et al. 2000). High school graduates in 1998 were more likely than their 1982 counterparts to take more advanced mathematics courses, such as algebra II, trigonometry, precalculus, and calculus. In science, the trend is similar. High school graduates in 1998 were more likely to take chemistry II or physics II and physics I and chemistry I (Wirt et al. 2000). Despite these encouraging signs, the experience is not reflected equally among racial/ethnic and income groups. In 1998, white and Asian/Pacific Islander high school graduates were usually more likely than black, Hispanic, and American Indian/Alaskan Native students to complete advanced academic-level mathematics and the highest-level science courses (Wirt et al. 2000). Students from low-income families were less likely to be enrolled in a college preparatory track through which they would be more likely to take such courses (U.S. Department of Education 1995).

Technology

Research suggests that student learning is enhanced by computers when the computer is used to teach discrete skills (President's Committee of Advisors on Science and Technology Panel on Educational Technology 1997). Computer availability and usage is increasing in the schools (Anderson and Ronnkvist 1999). In 1999, there was an average of 6 students for each computer, up from a 125 to 1 ratio in 1983 (Coley, Cradler, and Engel 1997; U.S. Department of Education 2000b). Internet access existed at 95 percent of public schools in 1999, up from 35 percent in 1994 (U.S. Department of Education 2000b). Internet access will most likely be used most if the computers are in instructional rooms. Over half (63 percent) of all instructional rooms (classrooms, computer or other labs, and library media centers) had access to the Internet in 1999, up from 3 percent five years before (U.S. Department of Education 2000b). For schools with high concentrations of poverty (more than 70 percent eligible for free or reduced-price lunch), 39 percent of all instructional rooms had Internet access compared with 62 to 74 percent for schools with lower concentrations of poverty (U.S. Department of Education 2000b).

Class Size

Researchers have found that greater gains in student achievement occur in classes with 13 to 20 students compared with larger classes, especially for disadvantaged and minority students (Krueger 1998; Mosteller, Light, and Sachs 1996; Robinson and Wittebols 1986). In 1998, the average public elementary school class had 23 students (Lewis et al. 1999). Large-scale efforts to reduce class size may result in negative consequences if, as was the case recently in California, large numbers of unqualified teachers are hired because there are not enough qualified teachers available to staff the smaller classes (Bohrnstedt and Stecher 1999).

Discipline

Researchers have found that a positive disciplinary climate is directly linked to student learning (Barton, Coley, and Wenglinsky 1998; Bryk, Lee, and Holland 1993; Chubb and Moe 1990). Research also suggests that the most effective policies to reduce the incidence of offenses in a school vary according to the targeted behavior. To reduce serious incidents, including drug offenses only a policy of severe punishment seems to be effective (Barton, Coley, and Wenglinsky 1998). Serious violent crime incidents, occurred in 10 percent of all public schools in 1996–97 (Kaufman et al. 1999). The level of school-related criminal behavior has changed little between 1976 and 1997, and no differences in victimization rates were found between white and black high school seniors in 1997 (Wirt et al. 1999). However, the percentage of middle and high school students who fear attack or other bodily harm while at school has been on the rise. In each year, a larger proportion of black and Hispanic students than white students feared attacks at school, and the percentage of black students who feared for their safety nearly doubled from 1989 through 1995 (Kaufman et al. 1999).

Academic Organization

Students learn more in schools that emphasize high academic expectations (Bryk, Lee, and Holland 1993; Chubb and Moe 1990), and academic expectations have been on the rise (Wirt et al. 1998). The percentage of public school districts with graduation requirements that meet or exceed the National Commission on Excellence in Education (NCEE) recommendations (four years of English, three years of mathematics, three years of science, three years of social studies, and a half year of computer science) increased from 12 to 20 percent between 1987–88 and 1993–94 (Wirt et al. 1998). A common criticism of the NCEE recommendations is that they only specify the number of courses to be taken, not their rigor. But there is evidence that increasing numbers of students have been enrolling in more difficult courses. From 1982 to 1998, there was an increase in the percentage of students enrolling in higher-level mathematics and science courses (Wirt et al. 2000).

SUMMARY

School quality needs to be defined, assessed, and monitored if we are to ensure the existence of quality schools (U.S. Department of Education 1991). This report highlights 13 indicators of school quality that recent research suggests may be related to student learning and identifies where and why more precise measures are needed. These indicators fall into three categories: the characteristics of teachers, the characteristics of classrooms, and the characteristics of schools as organizations. Research suggests that students learn more from teachers with high academic skills and teachers who teach subjects related to their undergraduate or graduate studies than they do from teachers with low academic skills and teachers who teach subjects unrelated to their training. In addition, students, on average, learn more from teachers with three or more years of teaching experience than they do from teachers with less experience. Though the research is less conclusive regarding professional development, experts agree that participation in high-quality professional development should lead to better teaching. At the level of the classroom, research suggests that students benefit from a focused and rigorous curriculum, time spent using computers, and being in smaller classes. We still need to learn more about the relationship between pedagogy and student learning. At the school level, a school's goals, leadership, faculty, discipline policy, and academic environment are all indicators of school quality. Student learning, however, is thought to occur primarily as a result of students' interaction with teachers, other students, and the curriculum, and the link between learning and these factors is not firmly established for all of these indicators.

Better measures are needed to accurately monitor the status of school quality, especially for indicators of pedagogy, leadership, goals, and professional community. Furthermore, certain important facets of professional development, course content, technology, academic environment, and discipline are missing. Finally, even when quality data are available, they lose their value if they are not appropriately defined and kept up to date. Moreover, even though experts would agree that certain indicators should enhance student learning, there is not always concrete statistical evidence to support their supposition; improving the data collected on the dimensions of schools thought to be associated with school quality should help us better understand the relationship of these indicators to student learning.

The findings documented in this report, like all research, are time sensitive and part of an iterative process. The status of schools as identified by indicators with quality data is changing rapidly and will need to be continually updated. As research on school effectiveness proceeds, indicators with only poor-quality data will need to be improved to understand the complete picture of school quality as recommended by the Special Study Panel on Education Indicators for the National Center for Education Statistics.

I. INDICATORS OF SCHOOL QUALITY

The nation's economic and social health depends on the quality of its schools. If students are not taught the values and social skills necessary to become good citizens and do not learn the academic skills necessary to be economically productive, then the schools have not succeeded in their mission. To ensure the existence of quality schools, school quality needs to be defined, assessed, and monitored (U.S. Department of Education 1991). This is one reason the congressionally mandated Special Study Panel on Education Indicators for the National Center for Education Statistics (NCES) called for indicator data to be regularly presented to the public "by interpretive reports that place data and analyses within the context of accessible written essays" (U.S. Department of Education 1991, p. 9). The panel called for these reports to focus on six issue areas: (1) school quality, (2) learner outcomes, (3) readiness for school, (4) societal support for learning, (5) education and economic productivity, and (6) equity (U.S. Department of Education 1991, p. 9). This report aims to help fulfill the panel's request in the first issue area, school quality.[1]

The report's three primary goals are:

1. Review the literature on school quality to help policymakers and researchers understand what is known about the characteristics of schools that are most likely related to student learning,
2. Identify where national indicator data are currently available and reliable, and
3. Assess the current status of our schools by examining and critiquing these national indicator data.

Student learning is, in part, a function of various characteristics of the schools and the process of schooling. Examining the characteristics of schools that are related to learning illuminates some of the reasons *why* students are, or are not, learning at optimum levels. In its 1991 report "Education Counts: An Indicator System to Monitor the Nation's Educational Health" (U.S. Department of Education 1991), the NCES panel called for reports that discuss indicators of the health of the nation's education system. Although it is less common to use indicators to measure the quality of schools than to gauge the health of the economy, there are compelling reasons to do so. For example, inflation, unemployment, growth rates, imports and exports, and wage growth are indicators that are looked at both separately and together to gauge the economy's strength. Unemployment might be low while inflation may be high. Each indicator reveals some important information about the economy and none alone tells the whole story. Similarly, school quality is simultaneously related to several characteristics of schools. It is a function of the training and talent of the teaching force, what goes on in the classrooms (including the size of the classes, the topics covered in them, and the pedagogical approaches used in them), and the overall culture and atmosphere within a school.

[1]This is the second interpretive report produced by NCES. The first report focused on education and the economy (Decker, Rice, and Moore 1997).

Defining school quality is the first step toward measuring and monitoring it. Both social and academic dimensions might be considered. The social includes the attitudes, ambitions, and mental well-being of students, while the academic dimension pertains to student learning. Both are important, but this report responds to the current national concern about academic quality by focusing solely on the school characteristics that have been shown to improve student learning.

Even though this report is concerned with what factors of schools are related to learning, this is not a report about test scores and the standards and accountability movement. This report is, however, still relevant to those interested in standards and accountability. It recognizes that some schools facing the same set of standards and accountability measures as other schools will do better in meeting them. This report tries to identify why this is the case. As the Action Statement from the 1999 National Education Summit noted, "Raising standards and developing tests may have been the easiest part of the journey; the more daunting task is ensuring that all students reach these standards" (1999 National Education Summit 1999). Because some schools do perform better than others, this report looks at the *components* that have been found to be linked to student learning and that may influence whether schools are successful in helping their students meet state and local standards.

Although most of the indicator data presented and critiqued in the report are national in scope, the report is intended to be useful to state and local policymakers and researchers. The data review will provide those interested in state and local data with an idea of what comparable data are needed to monitor state and local systems.

A. THE SCHOOL QUALITY LITERATURE

When thinking about what researchers have learned about school quality, it is useful to go back to the debate started by the "Equality of Educational Opportunity" report (Coleman et al. 1966), better known as the Coleman Report. This large national study sent shock waves through the education community when it concluded that measurable characteristics of teachers and schools are not significantly related to student achievement. Protracted debate has surrounded these findings, and numerous subsequent studies both support and refute the Coleman finding. The most comprehensive support, and the most widely cited, is Hanushek's (1986) review of the findings from 38 quantitative studies of the determinants of student achievement. These studies assessed the impact of the teacher/pupil ratio, teacher education, teacher experience, teacher salary, expenditures per pupil, administrative inputs, and facilities. Hanushek concludes that among these studies, "There appears to be no strong or systematic relationship between school expenditures and student performance" (Hanushek 1986, p. 1,162).

Many researchers do not accept Hanushek's conclusions. For example, Hedges, Laine, and Greenwald (Hedges, Laine, and Greenwald 1994) claim that his findings rest on inappropriate statistical methods and poor data. They conducted two separate studies to support their claim. In the first study, they reanalyzed studies Hanushek used and found a "positive relationship between dollars spent on education and output" that "is large enough to be of practical importance" (Hedges, Laine, and Greenwald 1994, p. 5).[2]

[2]The authors report that some of the studies used by Hanushek were excluded because they did not provide enough information to use in their analysis. The excluded studies were those that reported non-significant resource effects *and* did not report the sign of the coefficient.

To address the data quality limitations that they argue exist in Hanushek's work, Greenwald, Hedges, and Laine (1996) replicated the 1994 analysis using a more refined set of studies. The findings from this study bolster those from their earlier work.

Although Hanushek and Hedges, Laine, and Greenwald come to different conclusions, even Hanushek agrees that the data show that "teachers and schools differ dramatically in their effectiveness" (1986, p. 1,159). In more than 30 years of research and study on schooling and the educational process since the Coleman Report, a conclusive understanding of the definitive features of quality schools has yet to be found. However, it is apparent that no single factor guarantees school quality. As this report makes evident, school quality depends on multiple, interdependent elements. Furthermore, as discussed throughout this report, research also suggests that one of the main reasons some studies have failed to detect *why* schools differ in their effectiveness is because they use poor measures of school quality, and/or they entirely exclude certain elements essential to quality. This report identifies where more precise measures are needed and which phenomena have not been adequately addressed.

B. USING MORE PRECISE MEASURES

The need for more precise measures is evident when considering teacher qualifications. Numerous studies examine the relationship between teacher qualifications and student learning, but usually they measure whether teachers are certified, whether they have master's degrees, and how long they have taught. When teacher qualifications are measured in these ways, the evidence that they make a difference in school quality is inconsistent. This may be because these measures lack specificity. For example, knowing whether a teacher is certified or whether a teacher has a master's degree does not reveal anything about whether a teacher is well educated in the subject he or she is teaching. Many teachers with master's degrees are assigned to teach courses for which they have had no training. In addition, with the exception of brand new teachers, knowing how many years a teacher has taught tells us little about the teacher's skills. On average, researchers have not found a discernible difference between teachers with 5, 10, 15, or 20 years of experience. As will be discussed in the next chapter, recent research suggests that more precise measures of teacher training, experience, and skills are stronger predictors of student learning and school quality.

C. USING NEW MEASURES

Different aspects of the schooling process that may also be related to quality have traditionally been understudied in national data-collection efforts. For example, differences in student performance may well have to do with content coverage and instructional practice (how teachers deliver the content), but until very recently, data on classroom practices were in short supply in part because policymakers and researchers emphasized an input-output model when studying schools and because these attributes of the schooling process are harder to measure than traditional input indicators. As of the mid-1980s, reports began to highlight the importance of delving deeper into what schools actually do (see, for example, McKnight et al. 1987; Powell, Farrar, and Cohen 1985; U.S. Department of Education 1983). The late 1980s marked the first time that researchers and policy makers began to push for routine data collection on the schooling process (Murnane and Raizen 1988; Porter 1991; Shavelson et al. 1987; U.S.

Figure 1.1—School quality indicators and their relationship to student learning

Department of Education 1988) and national data are just now becoming available (see, for example, Hawkins, Stancavage, and Dossey 1998; Henke, Chen, and Goldman 1999).

D. IDENTIFYING INDICATORS OF SCHOOL QUALITY

This report highlights 13 indicators of school quality that recent research suggests are related to student learning and identifies where and why more precise measures are needed. These indicators are presented in Figure 1.1 and are categorized as the characteristics of teachers, the characteristics of classrooms, and the characteristics of schools as organizations. The figure illustrates that these school quality factors can affect student learning *both directly and indirectly*. For example, school context characteristics like school leadership can have an impact on teachers and what they are able to accomplish in the classroom, and this in turn may influence student learning. In addition, various teacher-level attributes can affect the quality of the classroom and in turn student learning. Traits at each of these levels can also directly affect student learning.

The report discusses each of these levels (school, teacher, classroom) in a separate chapter. The indicators in the chapter on teachers include the academic skills of teachers, teaching assignments, teacher experience, and exposure to professional development opportunities. Regarding classrooms, the indicators include content coverage, pedagogy, technology, and class size. The chapter on school organization addresses how a school's goals, leadership, faculty, discipline policy, and academic environment are indicators of school quality.

Each chapter begins with a review of the literature highlighting the particular aspects that have been found to be associated with student learning and merit tracking. After identifying why a particular aspect qualifies as an indicator, relevant data pertaining to the status of that indicator are presented. In cases where inadequate data exist, the report discusses why the data are limited and makes recommendations for improving these data.

II. TEACHERS

In the summer of 1998 when about half the prospective Massachusetts teachers failed the state's new licensing exam, a discussion about the implications ensued in the national media (Sandham 1998; Sterling 1998). What was this test measuring? Why did so many teachers fail? Are the prospective teachers who failed unqualified to teach? What does this say about the profession as a whole? Who should be allowed to teach in the public schools?

Some researchers suggest that school quality might be inseparable from teacher quality, implying that education reformers in Massachusetts and other states may need to use tough licensing exams or other teacher-related reforms to make meaningful changes in the schools. According to Hanushek (1992), "The estimated difference in annual achievement growth between having a good and having a bad teacher can be more than one grade-level equivalent in test performance" (p. 107). Rivkin, Hanushek, and Kain recently concluded in one study that teacher quality is the most important determinant of school quality:

> The issue of whether or not there is significant variation in the quality of schools has lingered, quite inappropriately, since the original Coleman Report. This analysis identifies large differences in the quality of schools in a way that rules out the possibility that they are driven by nonschool factors … we conclude that the most significant [source of achievement variation] is … teacher quality … (Rivkin, Hanushek, and Kain 1998, p. 32)

And yet, even though these researchers found that teacher quality is important, their data sets did not contain enough information to allow them to explain what exactly makes one teacher more or less effective than another (Rivkin, Hanushek, and Kain 1998). Other studies, as will be discussed below, suggest that to ensure excellence, teachers should:

- Have high academic skills,
- Be required to teach in the field in which they received their training,
- Have more than a few years of experience (to be most effective), and
- Participate in high-quality induction and professional development programs.

A. INDICATOR 1: THE ACADEMIC SKILLS OF TEACHERS

Many studies show that students learn more from teachers with strong academic skills than they do from teachers with weak academic skills (Ballou 1996; Ehrenberg and Brewer 1994; 1995; Ferguson 1991; Ferguson and Ladd 1996; Mosteller and Moynihan 1972). Because measures of teachers' academic skills are not routinely collected, the number of studies that look at this relationship is limited, and each uses a slightly different measurement method. The findings, however, are so consistent that there is broad agreement that teachers' academic skills are linked to student learning (Hanushek 1996; Hedges, Laine, and Greenwald 1994). This is not to say that academic skills perfectly predict how well a person will teach. Some educators argue that teacher quality has less to do with how well teachers perform on standardized tests than with how they perform in the classroom (Darling-Hammond 1998). In fact, classroom observation is the traditional way of assessing teacher quality. Obviously, several other traits not measured on standardized tests (such as interpersonal skills, public speaking skills, and enthusiasm for working with children) influence whether someone will be an effective teacher, but to date the only way these traits are systematically assessed is through formal classroom observation. Because these data are hard to quantify, most studies that have examined the link between teacher skills and student learning limit their definition of teacher skills to academic skills. We now will look at the findings from three of the most recent studies in this area.

Ehrenberg and Brewer (1994) investigated whether the quality of a teacher's undergraduate institution is related to student learning. Controlling for student and teacher background characteristics such as race/ethnicity and socioeconomic status, they found that the higher the quality of the institution a teacher attended, as measured by admission selectivity, the more students learned over the course of two years.[3] To the extent that the quality of a teacher's undergraduate institution is correlated with the academic skills of the teacher, this finding suggests that the more able teachers have students with higher scores.

Ferguson (1998) and Ferguson and Ladd (1996) used a more direct measure of the academic skills of teachers—their scores on standardized tests. These studies used state-specific data sets and, after controlling for several community and teacher characteristics such as race/ethnicity, found that higher teacher test scores are positively correlated with higher student test scores. Ferguson used Texas district-level data (from about 900 school districts) to measure the relationship between the average basic literacy skills of the teachers in a district and student learning gains over two years on mathematics tests. Ferguson reported that a one standard deviation change in the literacy skills of teachers would be associated with a 0.16 standard deviation increase in high school students' learning and a 0.18 standard deviation increase in elementary school students' learning.[4,5]

[3]The data come from the High School and Beyond (HS&B) study's 1984 supplementary teacher and administrative survey. This survey contains information about the undergraduate institutions teachers attended. The authors then linked these institutions to an admissions selectivity scale presented in Barron's (1984) and ranked the teachers' undergraduate institutions on a six-point scale ranging from most selective to least selective.

[4]When reviewing the research literature, this report will include an estimate, whenever possible, of how much of a boost in student learning or achievement is associated with a change in a particular component of school quality. These "effect size" estimates are presented as a fraction of a standard deviation so that they can be compared across studies. For example, if two different studies both find an effect size of 0.25, then it can be concluded that the size, or magnitude, of the effect on student learning is similar across studies. In education (and the behavioral sciences as a whole), when studies find effects, they tend to be modest in size, in the range of 0.10 or 0.20 of a standard deviation (Lipsey and Wilson 1993). As useful as effect size estimates are, there are unfortunately numerous important studies pertaining to school quality that do not provide enough information to allow for effect size estimates to be constructed. The relevance of all studies used in this report will be made clear in the chapters that follow, whether effect size estimates are presented or not.

[5]How can the magnitude of an effect size in standard deviation units be interpreted? If student test scores are normally distributed across a population and the average student scores better than 50 percent of that population, an effect size of 0.10 would boost the average student's score to be better than 54 percent of the population. An effect size of 0.25 would boost it to be better than 60 percent of the population, and an effect size of 0.50 would boost it to be better than 69 percent of the population.

In Alabama, Ferguson and Ladd had test scores from the teachers of almost 30,000 fourth-grade students in 690 schools. The scores were from the ACT exams the teachers took when they applied for college. Over the course of one year, Ferguson and Ladd found that a one standard deviation difference in a school's distribution of teacher ACT scores was associated with a 0.10 of a standard deviation change in the distribution of that school's fourth-grade reading test scores.

What cumulative impact will raising the overall academic caliber of teachers have on student learning from grade 1 through grade 12? Unfortunately, this is currently unknown. Even though the effect sizes reported in these two studies are modest, they show impacts only over a one- and two-year period. Do students who are annually taught by higher-caliber teachers receive persistent advantages (beyond two years) compared with their counterparts in lower-caliber teachers' classrooms? Are these gains of the same magnitude year after year? If there are annual gains, the effect sizes presented above may greatly underestimate the benefit students would receive throughout their schooling from being taught by more academically able teachers.

Given that students learn more from teachers with strong academic skills than they do from teachers with weak academic skills, it would be useful to monitor the academic strength of the teaching force. How do the academic skills of teachers compare with other professionals? Is the academic talent of teachers distributed evenly among different types of schools?

Several studies show that over the past three decades, teachers with low academic skills have been entering the profession in much higher numbers than teachers with high academic skills (Ballou 1996; Gitomer, Latham, and Ziomek 1999; Henke, Chen, and Geis 2000; Henke, Geis, and Giambattista 1996; Murnane et al. 1991; Vance and Schlechty 1982).[6]

Murnane et al. (1991) found that entering teacher IQ scores declined from the 1960s through the 1980s. In 1967, graduates with IQ scores of 100 and 130 were equally likely to become teachers, but by 1980, the ratio was 4 to 1.[7] In other words, in 1967, for every four graduates with an IQ of 100 who entered the teaching profession, there were four graduates with an IQ of 130 who entered the profession. In 1980, for every four graduates with an IQ of 100 who entered the profession, there was only one graduate with an IQ of 130. Vance and Schlechty found that in the 1970s teaching attracted and retained a disproportionately high share of college graduates with low SAT scores and failed to attract and retain those with high SAT scores (Vance and Schlechty 1982).[8]

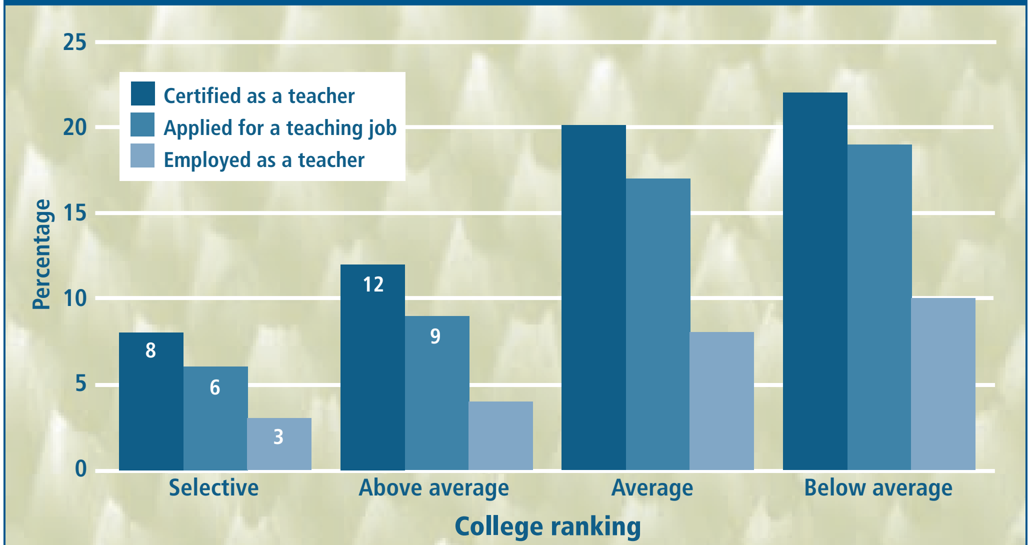
Evidence suggests that these trends have persisted into the 1990s. Ballou (1996) found that the higher the quality of the undergraduate institution attended, as measured by

[6]Two studies (Bruschi and Coley 1999; Rollefson and Smith 1997) using one data source, the National Adult Literacy Survey (NALS), found that in 1992, the teachers in the U.S. schools had literacy skills similar to those of professionals in several other occupations for which a bachelor's degree is a prerequisite. These professionals included physicians, engineers, postsecondary teachers, writers, and artists. The NALS data differ from the data used in these other studies in that they pertain to (Bruschi and Coley 1999; Rollefson and Smith 1997) literacy skills, as opposed to a more general set of academic skills, and to the skills of existing teachers, not the skills of new entrants.

[7]This study used the National Longitudinal Surveys of Labor Market Experience. These surveys contain nationally representative information on individual characteristics, education, employment, and teaching status.

[8]This study used the National Longitudinal Study of 1972 high school seniors.

Figure 2.1—Percentage of teachers at various stages of new teacher recruitment, by college ranking: 1976–1991



SOURCE: Dale Ballou, "Do Public Schools Hire the Best Applicants?," *The Quarterly Journal of Economics* (February 1996): 97–133, Survey of Recent College Graduates, pooled data from six administrations of the survey that occurred between 1976 and 1991.

the Barron's admissions selectivity scale, the *less* likely a student is to prepare to become a teacher and enter the teaching profession. Ballou used the Surveys of Recent College Graduates to sort students by the selectivity of their undergraduate institutions (the ratings range from selective to below average) and then examined the rate at which students at these different types of institutions took the courses necessary to become certified teachers, applied for a teaching job, and actually became teachers. Figure 2.1 shows that the less selective the college, the more likely that students at that college will prepare for and enter the teaching profession.[9] Ballou concluded, "Thus, certification, application, and employment levels all rise monotonically as college quality declines" (1996, p. 103).

Ballou's study was not the only study to use 1990s data to suggest that the teaching profession attracts those with lower academic skills. The Educational Testing Service (ETS) found that this was true for most of the prospective teachers taking the Praxis II exam between 1994 and 1997 (Gitomer, Latham, and Ziomek 1999). When comparing the average SAT scores for teacher candidates passing the Praxis II exam with the average SAT score for all college graduates, ETS concluded that elementary education candidates, the largest single group of prospective teachers, have much lower math and verbal scores. The pattern in other content areas for teacher candidates was less consistent. The average math SAT score for those passing the Praxis II exam and seeking licensure in physical education, special education, art and music, social studies, English, or foreign language was lower than the average math score for all college graduates. Those seeking to teach science and math, however, had higher average math

[9]This figure includes pooled data from the six administrations of the Surveys of Recent College Graduates that occurred between 1976 and 1991.

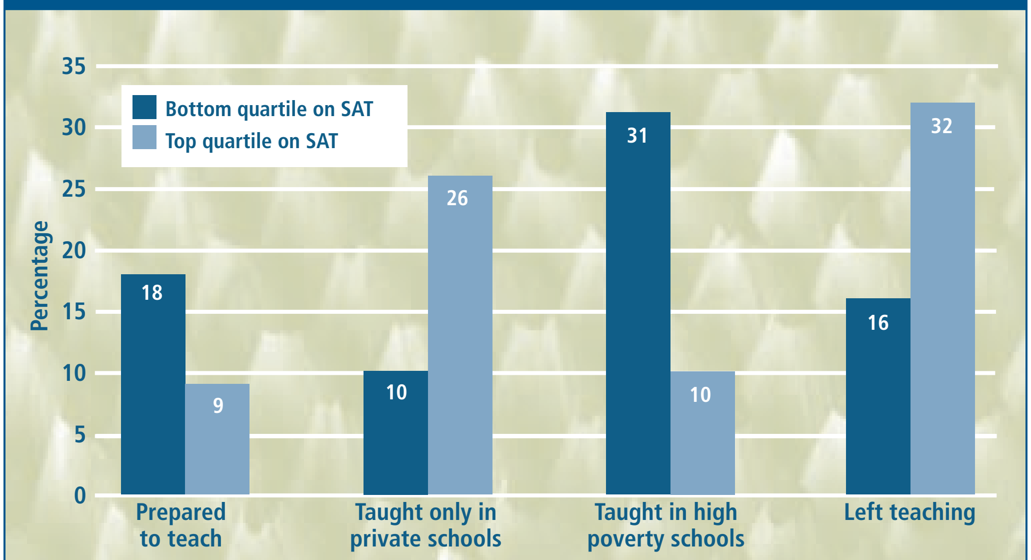
scores. The average verbal SAT scores were more encouraging. The scores of mathematics, social studies, foreign language, science, and English candidates who passed the Praxis II exam were as high or higher than the average verbal SAT score for all college graduates. Physical education, special education, and art and music teachers scored below the average.

A limitation of this analysis is that it provides data only on candidates, not actual teachers. As Ballou's data in Figure 2.1 show, there are large drop-offs in the pipeline. For example, while 20 percent of students from average colleges became certified to teach, 17 percent applied for teaching jobs, and 8 percent actually became employed as teachers. Given such large drop-offs in the pipeline, we cannot just assume that those who pass the Praxis examination have the same characteristics as those who actually end up teaching.

Recent studies, using data from the 1993 Baccalaureate and Beyond Longitudinal Study, provide a more comprehensive picture of the pipeline from preparation to employment (Henke, Chen, and Geis 2000; Henke, Geis, and Giambattista 1996). These studies found that the college entrance examination scores of the 1992–93 college graduates in the teaching pipeline (defined by NCES as students who had prepared to teach, who were teaching, or who were considering teaching) were lower than those students who were not in the pipeline. "At each step toward a long-term career in teaching, those who were more inclined to teach scored less well than those less inclined to teach" (1996, p. 21). For example, as shown in Figure 2.2, by 1997 the 1992–93 college graduates in this study with the highest college entrance examination scores were consistently less likely than their peers with lower scores to prepare to teach, and when they did teach, they were less likely to teach students from disadvantaged backgrounds:

- Graduates whose college entrance examination scores were in the top quartile were *half* as likely as those in the bottom quartile to prepare to teach (9 versus 18 percent).

Figure 2.2 Percentage of 1992–1993 college graduates who prepared to teach, where they taught, and who left teaching by SAT scores: 1997



SOURCE: U.S. Department of Education, National Center for Education Statistics, 1993 Baccalaureate and Beyond Longitudinal Study, Second Follow-up (B&B; 93/97), Data Analysis System.

- Teachers in the top quartile were more than *twice* as likely as teachers in the bottom quartile to teach in private schools (26 versus 10 percent).
- Teachers in the top quartile were at least *one-third* as likely as teachers in the bottom quartile to teach in high-poverty schools (10 versus 31 percent).
- Graduates in the top quartile who did teach were *twice* as likely as those in the bottom quartile to leave the profession within less than four years (32 versus 16 percent) (Henke, Chen, and Geis 2000).

These studies show a consistent trend and suggest that there is a need to monitor closely the supply and distribution of teacher academic skills. Unfortunately, the national data on teacher academic skills currently available are limited by their lack of specificity, timeliness, generalizability, and ability to link to student performance. The Survey of Recent College Graduates ascertains the academic quality of the undergraduate institution a person attended, but it does not reveal whether the person was in the top or bottom of the academic distribution at that institution. The National Adult Literacy Study and the Baccalaureate and Beyond Longitudinal Study provide information about how teachers' academic skills compare with those of other professionals, but neither study allows for a link to student performance. While some currently available data give a more direct measure of an individual teacher's academic ability and can be linked to student test scores (Ferguson 1998; Ferguson and Ladd 1996), the data are not collected routinely and are limited to a few states. Better nationally representative data are needed to gauge several aspects: how the academic caliber of teachers compares with that of other professionals; how the existing teaching talent is distributed throughout the country; and how teachers' academic skills have a cumulative impact on student academic performance.

B. INDICATOR 2: TEACHER ASSIGNMENT

Many teachers are currently teaching courses they were not trained to teach, and this appears to affect student achievement adversely (Darling-Hammond 2000; Goldhaber and Brewer 1997; Monk and King 1994). Though several studies show mixed results concerning the relationship between teacher degree and student test scores, most of these studies simply assess whether a teacher has a master's degree (for a review of the results from these studies see Greenwald, Hedges, and Laine 1996; Hanushek 1989) and do not identify the subject in which the degree was received or the type of training a teacher received.

Goldhaber and Brewer (1997), Darling-Hammond (2000), and Monk and King (1994) found that subject matter preparation is related to student achievement even after controlling for relevant teacher and student background and contextual variables such as race/ethnicity and socioeconomic status. Goldhaber and Brewer (1997) confirmed this significant relationship in mathematics and science but found no effect in English and history.[10] Teachers with bachelor's and master's degrees in mathematics are associated with higher student mathematics test scores. Teachers with bachelor's degrees in science

[10]This study uses data from the National Education Longitudinal Study of 1988 to look at the relationship between teacher characteristics and student achievement scores of 10th grade students.

are associated with higher student science scores. The effect size in both instances is about 0.10 of a standard deviation.[11]

Monk and King (1994) looked at the relationship between the preparation of mathematics and science teachers and student learning. Using nationally representative data,[12] they measured preparation by counting the number of graduate and undergraduate courses teachers took in their field. Monk and King found that, in some instances, high school students' mathematics and science test scores are associated with the subject-matter preparation of their teachers. (They did not examine English or history test scores.) The results for mathematics, however, are stronger and larger when they include the cumulative mathematics preparation of all the mathematics teachers that students had in both their sophomore and junior years in high school. The students who scored below the median on a pretest appeared to reap the most benefits. Cumulative effects were not found in science.

Darling-Hammond (2000) conducted a state-level analysis examining the relationship between teacher preparation and 4th and 8th grade student achievement on the National Assessment of Educational Progress math and reading exams. After controlling for the percentage of students in poverty, the percentage who have limited English proficiency, average class size, and the percentage of teachers with master's degrees, she found that "the proportion of well-qualified teachers (those holding state certification and the equivalent of a major in the field taught) is by far the most important determinant of student achievement" (p. 27).

Given the apparent benefits students receive from being taught by well-qualified teachers, it is worth assessing the extent to which students are taught by teachers who are teaching without the proper qualifications. A frequently cited measure of whether a teacher is unqualified is one that determines whether a teacher is teaching out-of-field or teaching subjects that he or she was not trained to teach (Ingersoll 1999). Because this occurs mainly in the secondary and not the elementary grades (Bobbitt and McMillen 1994; Henke et al. 1997), this discussion focuses on the secondary level. There are two steps to defining out-of-field teaching: defining field of expertise and determining the number of courses taught by those without the proper qualifications or training. Some believe a secondary teacher's field is defined by the teacher's undergraduate or graduate major or minor. If she majored or minored in mathematics, her field is mathematics. Others argue that field should be defined as the subject in which the teacher is state certified, independent of her major or minor. Still others think that a teacher's field should be defined by the combination of major and minor *and* certification. A math teacher, for example, would have to have both majored in mathematics and been certified to teach in mathematics. Several reports present data pertaining to each of these definitions (Bobbitt and McMillen 1994; Henke et al. 1997; Ingersoll 1999), but there is some consensus that having an undergraduate or graduate major or minor is a minimal requirement (Ingersoll 1999), and that definition is used in the following discussion.

[11]The effect size estimates presented here differ slightly from those presented in the paper cited because the estimates in the paper were incorrect (personal communication with Goldhaber, March 1999). Both sets of estimates were calculated using the coefficients presented in Table 3 of the paper. However, the estimates in their paper were calculated using the coefficients in columns one and two for mathematics and three and four for science. Because columns one and three present misspecified models, the effect sizes should have been calculated using column two for mathematics and column four for science.

[12]The Longitudinal Study of American Youth.

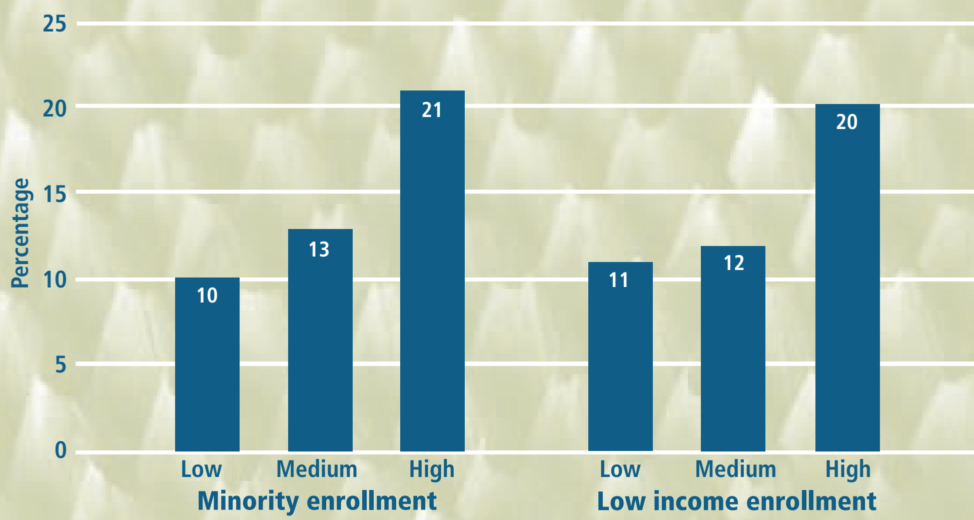
After defining in-field and out-of-field, researchers can estimate by field the extent of out-of-field teaching. There are three ways to measure and report out-of-field teaching. First, researchers have commonly reported these numbers at the *teacher level* (Bobbitt and McMillen 1994; Henke et al. 1997; Ingersoll 1999; Lewis et al. 1999). They report, for example, the percentage of teachers teaching mathematics who do not have the proper training ("the percent of teachers teaching math out-of-field"). However, because most teachers who teach out-of-field do not teach *all* of their courses out-of-field, this approach either underestimates or overestimates the problem. Researchers have used two imperfect teacher-level definitions of out-of-field teaching, neither of which can accurately account for the fact that teachers are often only partially out-of-field: the "any-mismatch" approach and the "main-assignment" approach (Bobbitt and McMillen 1994; Henke et al. 1997; Ingersoll 1999). The "any-mismatch" approach labels teachers as out-of-field if they are teaching *at least one* course that does not match their field (however field is defined). The "main-assignment" approach labels teachers as out-of-field only if *most* of the courses they teach do not match their field. For example, a teacher who is certified in social studies and teaches four of her five courses in social studies and one of her five courses in math would be considered teaching out-of-field in math when using the first approach ("any-mismatch") but not the second approach ("main-assignment"). Consequently, the "main-assignment" approach underestimates the magnitude of the "out-of-field" phenomenon (Bobbitt and McMillen 1994; Lewis et al. 1999) because it counts this teacher as teaching in-field even though she is teaching math out-of-field. Conversely, the "any-mismatch" approach overestimates the magnitude of the problem because it counts this teacher as an out-of-field math teacher even though she is teaching only one math course (not her entire course load) out-of-field. Precisely because teachers usually do not teach all of their courses out-of-field, it is not optimal to assess the *percentage of teachers* teaching out-of-field in a given subject.

The other two approaches come closer to assessing the true magnitude of the out-of-field phenomenon. It is more informative to assess what *percentage of courses* in given subjects are taught by out-of-field teachers and, because not all classes have the same numbers of students, the *percentage of students* in given subjects taught by out-of-field teachers. These two measures identify for policy makers the extent of the qualified teacher shortfall and will pinpoint the percentage of students affected by the problem.

The percentage-of-courses measure has not been used in prior analyses of national data. The percentage-of-student measure has been generated (Bobbitt and McMillen 1994; Ingersoll 1999), but unfortunately it has not been generated using the most recently available data (the 1993–94 Schools and Staffing Survey and the 1999 Teacher Quality Survey) . The most recently available student-level analyses provides data from the 1990–91 school year (Bobbitt and McMillen 1994; Ingersoll 1999). This analysis shows that 14 percent of social studies students, 23 percent of English/language arts students, 18 percent of science students, and 30 percent of mathematics students in public secondary schools (grades 7 through 12) were taught by teachers who did not major or minor in these fields (Bobbitt and McMillen 1994).

Unlike some of the other indicators discussed in this report (such as indicators of pedagogy and school leadership discussed in subsequent chapters), measuring out-of-field teaching is relatively straightforward. Even though there are various ways to define "qualified," the types of survey questions needed to assess training and certification are known. And even though there are various ways to count how many courses are taught by unqualified teachers, there are meaningful measures that can be constructed. As new

Figure 2.3 Percentage of teachers with three or fewer years of experience by level of minority and low income enrollment: 1998



NOTE: Low, medium and high are defined as the schools in the bottom quartile, the middle two quartiles, and the top quartile, respectively. Low income is defined as the percent of students eligible for free or reduced price lunch.

SOURCE: Fast Response Survey System: Teacher Survey on Professional Development and Training. 1998.

data become available from the 2000–01 Schools and Staffing Survey and the 2000 Teacher Quality Survey, student and course-level estimates will be the most meaningful and precise estimates of the extent of in-field and out-of-field teaching.

C. INDICATOR 3: TEACHER EXPERIENCE

Studies suggest that students learn more from experienced teachers than they do from less experienced teachers. Murnane and Phillips (1981) reported that in a large city in the Midwest, after controlling for other student and teacher characteristics such as race/ethnicity and socioeconomic status, children taught by a teacher with five years of experience make three to four months' more progress in reading skills during a school year than do children taught by a first-year teacher. A more recent study conducted by Rivkin, Hanushek and Kain (1998) found that 4th, 5th, and 6th grade students in more experienced teachers' classrooms in Texas over the course of one year gained about 0.10 of a standard deviation in reading and math compared with their peers in classrooms where teachers had less than two years of experience. The benefits of experience, however, appear to level off after 5 years, and there are no noticeable differences, for example, in the effectiveness of a teacher with 5 years of experience versus a teacher with 10 years of experience (Darling-Hammond 2000). However, teachers with 5 or 10 years of experience are more effective than new teachers.

Though it is impossible to limit the teaching force only to experienced teachers, the effects of new teachers may be diffused and reduced if new teachers are evenly distributed among the schools, and proper assistance is given to new teachers.

As of 1998, teachers with three or fewer years of experience were not spread evenly among different types of schools. Figure 2.3 shows that the highest-poverty schools and schools with the highest concentrations of minority students (those in the top quartile) have a higher proportion of inexperienced teachers than schools with lower levels of poverty and lower numbers of minority students (those in the three other quartiles). The highest-poverty schools and schools with the highest concentrations of minority students had nearly *double* the proportion of inexperienced teachers as schools with the lowest poverty (20 versus 11 percent) and lowest concentration of minority students (21 versus 10 percent). One likely cause for this overrepresentation of inexperienced teachers is that teacher attrition disproportionately affects high-poverty schools (Henke et al. 1997).

D. INDICATOR 4: PROFESSIONAL DEVELOPMENT

The quality of the teaching force may depend on the opportunities for development presented to those already teaching because entering teachers make up a minority of the teaching corps. Even though experts would likely agree that professional development should enhance student learning, there is no concrete statistical evidence of an association. This lack of statistical evidence may be because the quality of the data pertaining to professional development needs to be improved to understand more about its relationship with student learning.

In the 1980s and 1990s, large numbers of teachers left the profession within the first few years of entering it. For example, between the 1993–94 and 1994–95 school years, the most recent years in which national attrition data exist, 17 percent of teachers with three or fewer years of experience left the profession. Nine percent left after teaching for less than one year. And, as noted above, a disproportionately high share left high-poverty schools.

Further studies using both state and national data have shown that the most academically able teachers are the most likely to leave the profession in the first few years (Henke, Chen, and Geis 2000; Heyns 1988; Murnane and Olsen 1990; Vance and Schlechty 1982). This compounds the problem identified above that the most academically talented may be the least likely to enter the profession in the first place (Ballou 1996; Haney, Madaus, and Kreitzer 1987; Henke, Chen, and Geis 2000; Heyns 1988; Murnane and Olsen 1990; Vance and Schlechty 1982). In addition, as discussed in the school chapter below, high teacher attrition may negatively affect a school's professional community and student learning.

In several administrations of the Schools and Staffing Survey (1988–89, 1991–92, 1994–95), teachers who reported "dissatisfaction with teaching as a career" as one of the three main reasons for leaving teaching were asked what specifically they were dissatisfied with. Among the top concerns cited in each survey were "inadequate support from administration," "poor student motivation to learn," and "student discipline problems" (Whitener et al. 1997).

To keep young teachers committed to the profession and to help them learn the trade, the National Commission on Teaching and America's Future (1996) recommends that schools institute induction programs. The commission suggests that these programs should be modeled on the residency programs used in medicine and should include the pairing of beginning teachers with skilled mentors. Formal induction programs appear to be on the rise. For example, teachers with three or fewer years of experience were more likely to have reported participating in an induction program in 1998–99 than in

1993–94 (65 versus 59 percent) (Lewis et al. 1999). Little is known about the form these programs take, whether they will help novice teachers teach better, or whether they will stem attrition.

Veteran teachers also have professional development needs. Several reports have asserted that teachers will perform better if they are given opportunities to sharpen their skills and keep abreast of advances in their field (Henke, Chen, and Geis 2000; National Commission on Teaching and America's Future 1996), though a comprehensive assessment of the availability of such learning opportunities and their impact on teachers and students has yet to be done (Mullens et al. 1996; Smylie 1996).

Nevertheless, several reform initiatives have noted that "professional development" (PD) should play a central role in improving the schools (National Commission on Teaching and America's Future 1996; National Education Goals Panel 1995; National Foundation for the Improvement of Education 1996). The National Education Goals Panel endorsed high-quality professional development in 1994 by setting the following goal: "the nation's teaching force will have access to programs for the continued improvement of their professional skills and the opportunity to acquire the knowledge and skills needed to instruct and prepare all American students for the next century" (National Education Goals Panel 1995, p. 93).

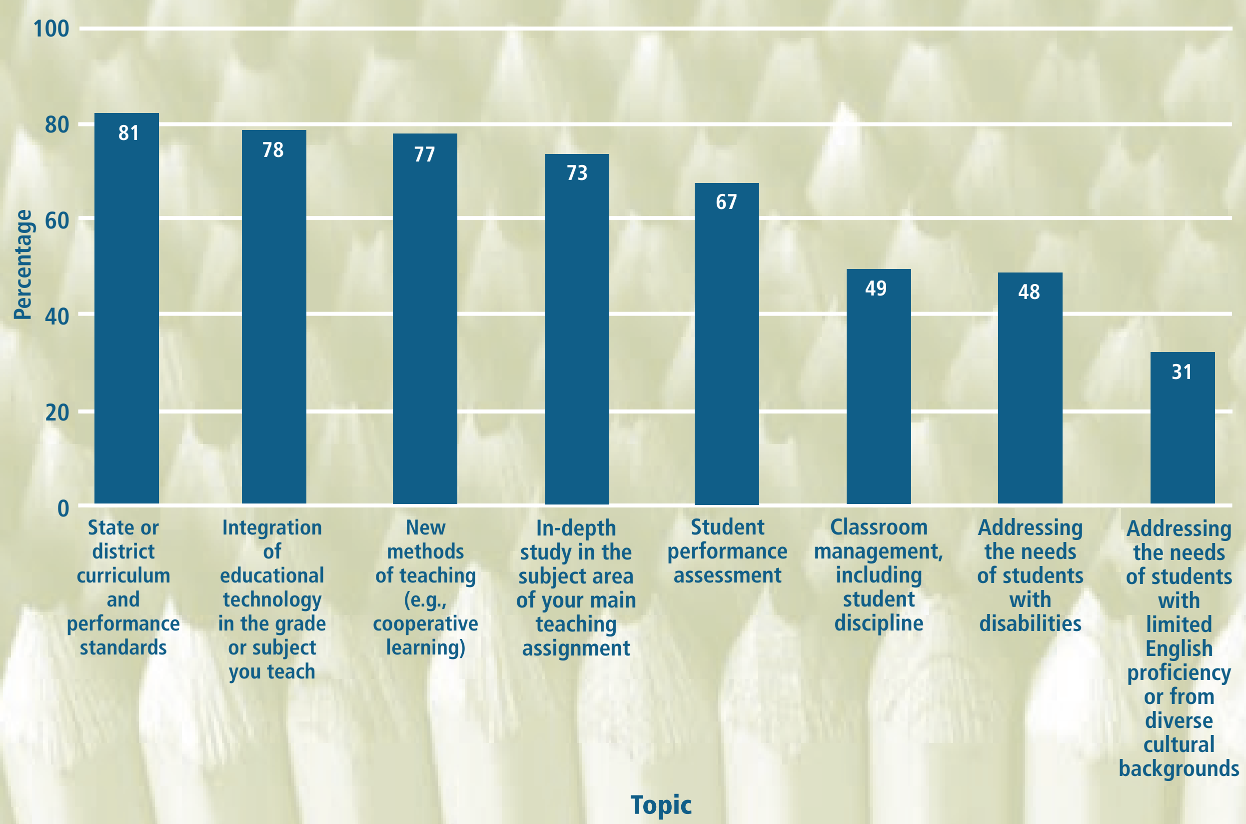
Part of the reason for this support is that a high percentage of the teaching force consists of teachers who received their initial training more than 20 years ago. In 1998, 64 percent of public school teachers had 10 or more years of experience, and 39 percent had 20 or more (Lewis et al. 1999). In other words, without formal PD initiatives, a substantial number of teachers might be uninformed about key advances that have occurred in the field of education since they received their initial training. PD advocates believe that the overall quality of the nation's teachers depends on teachers being given the opportunity to learn about new theories of teaching and learning, changes in the student population, and how to use new technologies (such as computers and the Internet) in their classrooms (Choy and Ross 1998; National Education Goals Panel 1995; National Foundation for the Improvement of Education 1996).

There is broad consensus about the elements that constitute an effective professional development program (CPRE Policy Brief 1995; National Commission on Teaching and America's Future 1996; National Foundation for the Improvement of Education 1996; U.S. Department of Education 1999a). The National Education Goals Panel cited several of these elements in its list of "principles of high quality professional development programs" (Goals 2000 1999, p. 2). Successful programs:

- Focus on individual, collegial, and organizational improvement,
- Promote continuous inquiry and improvement embedded in the daily life of schools,
- Are planned collaboratively by those who will participate in and facilitate that development,
- Require substantial time and other resources, and
- Are driven by a coherent long-term plan.

In addition to these five principles, research by Cohen and Hill (2000) suggests that professional development activities that are tightly linked to well-defined instructional goals result in improved teaching. To date, the degree to which PD activities across the

Figure 2.4 Percentage of full-time public school teachers who participated in professional development activities in the last 12 months that focused on various topics: 1998



SOURCE: U.S. Department of Education, National Center for Education Statistics, *Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers,* Fast Response Survey System. 1999.

country embrace these principles is unknown. Current data tell us mostly about the prevalence of PD, but not much about its structure and quality.

What is known is that PD is ubiquitous in public schools. In 1998, 99 percent of the nation's public school teachers had participated within the previous 12 months in at least one of the eight PD activities listed in Figure 2.4. NCES concluded that these patterns indicated that teachers were engaged in professional development activities consistent with guidelines stipulated in recent education reforms (Lewis et al. 1999). However, as Figure 2.5 illustrates, the majority of teachers participated in these activities from one to eight hours, or for no more than one day. Thus, most teachers are not engaged in PD on particular topics for substantial amounts of time.

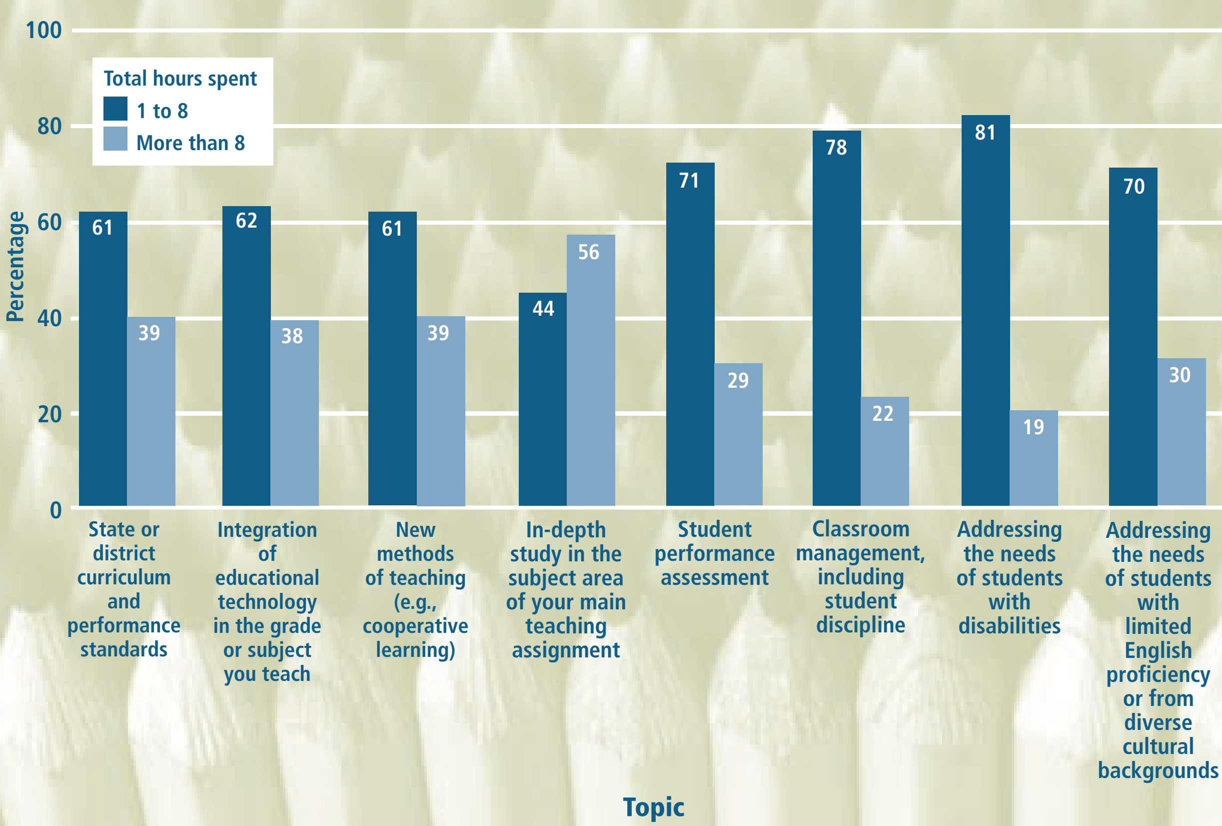
Beyond topic coverage and time spent within topic areas, national surveys reveal little about the caliber of PD activities. Additional data are needed to assess whether any of these teachers, even those who are participating in PD activities that last longer than eight hours, are engaged in activities that are part of a larger process of staff renewal and enhancement consistent with the broadly accepted principles listed above.

E. SUMMARY

If, as Hanushek (1992) has suggested, teacher quality can translate into a difference in annual student achievement growth of more than one grade level, then teacher quality may be among the most important issues to consider when thinking about school quality. Numerous studies examine the relationship between teacher qualifications and student learning, but teacher qualifications are most commonly measured in these studies by looking at whether teachers are certified and whether they have master's degrees. These measures of teacher qualifications do not appear to be related to school quality, perhaps because they lack specificity. Current research using more precise measures suggests that teacher quality with respect to student outcomes might be improved if teachers' academic skills are improved, if more teachers teach in the field in which they received their training, and possibly if teachers participate in high-quality professional development activities related to content.

National data show that poor children receive less than their fair share of high-quality teachers. The academic skills of incoming teachers are relatively weak compared with the average college student; many teachers, especially math teachers, are teaching subjects they were not trained in; and many teachers do not experience sustained professional develop-

Figure 2.5 Percentage of full-time public school teachers indicating the number of hours spent in professional development activities on various topics in the last 12 months: 1998



SOURCE: U.S. Department of Education, National Center for Education Statistics, *Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers,* Fast Response Survey System. 1999.

ment experiences to help them grow and learn on the job. And quality among the current teaching corps is not evenly distributed throughout the nation. High-poverty schools and high-minority schools have a disproportionately high share of inexperienced teachers relative to low-poverty and low-minority schools; and high-poverty schools have a disproportionately high share of academically weak teachers relative to low-poverty schools.

III. CLASSROOMS

Within the classroom, the curriculum is the cornerstone of the academic experience. According to researchers who worked on the Second International Mathematics Study, the curriculum has three levels (McKnight et al. 1987):

- The *intended* curriculum, which is defined by officials at the state and local levels;
- The *implemented* curriculum, which is how teachers translate the intended curriculum into practice; and
- The *attained* curriculum, which is what students learn as represented by their scores on standardized tests.

The intended curriculum consists of the topics that teachers are expected to teach. Curriculum topics are often prescribed by content standards set by states and local school districts; those standards are frequently influenced by national organizations such as the National Council for Teachers of Mathematics and the National Academy of Sciences. The intended curriculum is represented most directly by the curriculum materials provided to teachers: curriculum guides, textbooks, videos, computer software, and other curriculum-related items. Because the intended curriculum is, at best, a set of goals for what teachers and students address in the classroom, the intended curriculum has limited value to those interested in knowing how those goals are implemented and how they relate to student learning.

The implemented curriculum includes that part of the intended curriculum that is actually taught, the instructional practices and tools employed, and the conditions under which those things happen. It has direct implications for how course content is covered, how much time is available for other topics, and, ultimately, how much and how well students might learn. Aspects of the implemented curriculum that might reflect school quality include not only the content of the implemented curriculum (the curriculum topics actually addressed in the classroom), but also the pedagogy (or instructional practices), the materials and equipment (such as technology), and the conditions under which the curriculum is implemented (such as the number of students in the class). Furthermore, each must be related to student learning.

The attained curriculum is the final product of that part of the intended curriculum that was implemented. These three linked levels of curriculum become a classroom input-output model where the intended curriculum represents the inputs, the attained curriculum equals the outputs, and the implemented curriculum is the black box of the classroom and what happens inside it. In this model, the implemented curriculum has the greatest potential variation and is projected to be the only direct effect on student learning.

Inadequate data, however, make it difficult to measure the impact of the implemented curriculum on student learning. Before the mid-1980s, nationally representative studies (such as High School and Beyond and the National Assessment of Educational Progress) limited their data collection to elements of the intended curriculum and other inputs, such as per-pupil spending, course titles, and teacher salaries. Then several reports highlighted the importance of determining what happens in classrooms by studying elements of the implemented curriculum, such as curricular content, instructional strategies, and organization (McKnight et al. 1987; Powell, Farrar, and Cohen

1985; U.S. Department of Education 1983). Therefore, in the late 1980s, researchers and policymakers began for the first time to push for the regular collection of data that provide information on the schooling process (Murnane and Raizen 1988; Shavelson et al. 1987; U.S. Department of Education 1988).

Policymakers and researchers believe that four elements of the implemented curriculum hold the most promise for gaining insight into the relationship between classrooms and student learning: course content, instructional practices or style, classroom technology, and class size. The remainder of this chapter discusses findings from research and issues related to collecting data on these aspects of classrooms.

A. INDICATOR 5: COURSE CONTENT

Concerns about content in both the intended and implemented curriculum and how a lack of focus in the U.S. curriculum affects student achievement have appeared in influential studies since the early 1980s. In 1983, the National Commission on Excellence in Education (NCEE) concluded:

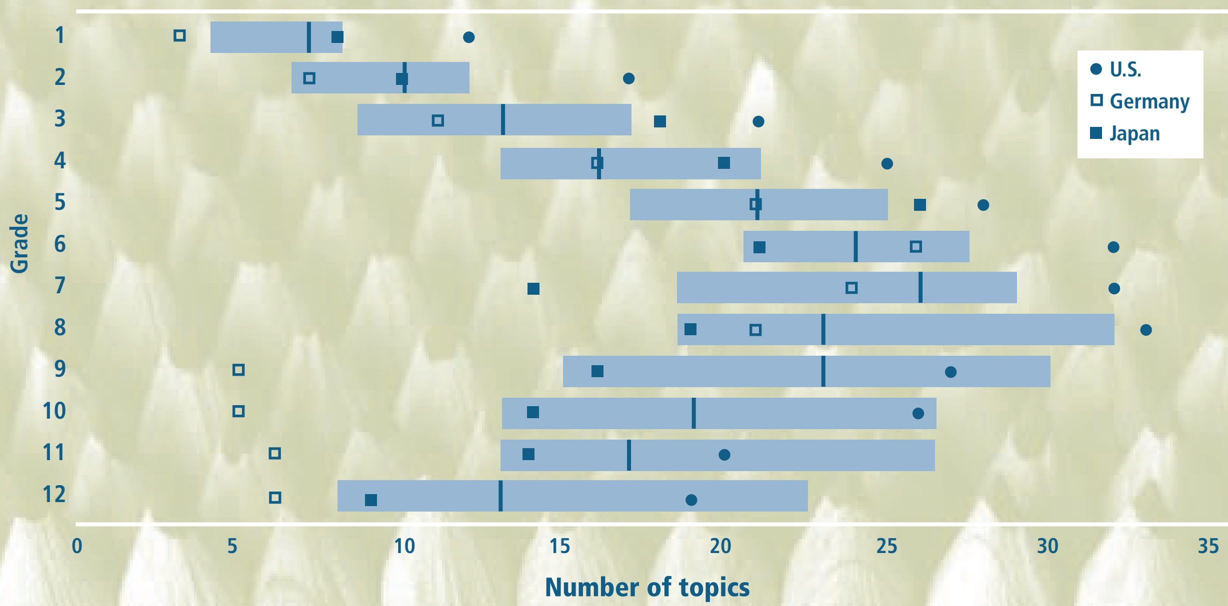
> Secondary school curricula have been homogenized, diluted, and diffused to the point that they no longer have a central purpose. In effect, we have a cafeteria-style curriculum in which the appetizers and desserts can easily be mistaken for the main courses. Students have migrated from vocational and college preparatory programs to "general track" courses in large numbers. The proportion of students taking a general program of study has increased from 12 percent in 1964 to 42 percent in 1979. This curricular smorgasbord, combined with extensive student choice, explains a great deal about where we find ourselves today (U.S. Department of Education 1983, p. 18).

The NCEE recommended that the curriculum be brought under control by requiring high school students to take the "Five New Basics:" four years of English, three years of mathematics, three years of science, three years of social studies, and a half year of computer science. While these recommendations were advanced for the times, they were stated only in terms of numbers of courses, not about particular courses or the rigor or topics within a course. The authors of *The Shopping Mall High School* seemed to note this distinction: "High schools seem unlikely to make marked improvement, especially for the many students and teachers now drifting around … until there is a much clearer sense of what is most important to teach and learn, and why, and how it can best be done" (Powell, Farrar, and Cohen 1985, p. 306).

Does a lack of content focus result in poor performance? "Yes," according to an analysis of results from the Second International Mathematics Study (SIMS) in the early 1980s. An analysis of the curriculum guides from the Third International Mathematics and Science Study (TIMSS) countries in the early 1990s also provides some useful data, though links to achievement are not clear-cut. In both studies, U.S. 8th and 12th graders scored low on international comparisons in math and science.[13] In 1987, McKnight et al. directed the blame at a diffuse curriculum:

[13]Research scientists suggest there are clear limitations in using SIMS and TIMSS data to make international comparisons in that neither study design adequately controlled for substantial differences in student selectivity, curriculum emphasis, and the proportion of low-income students in the test-taking populations (Rotberg 1998).

Figure 3.1 Number of mathematics topics intended to be covered in 36 countries



NOTE: The light blue bars show how many mathematics topics were intended to be covered at each grade in the TIMSS countries. The bars extend from the 25th percentile to the 75th percentile among countries. The blue line indicates the median number of topics at each grade.

SOURCE: William H. Schmidt, Curtis C. McKnight, and Senta A. Raizen, *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education,* Third International Mathematics and Science Study, 1992–93.

> The culprit that seems to be central to the problems of school mathematics is the curriculum … Content is spread throughout the curriculum in a way that leads to very few topics being intensely pursued. Goals and expectations for learning are diffuse and unfocused. Content and goals linger from year to year so that curricula are driven and shaped by still-unmastered mathematics content begun years before [p. 9].

A decade later, TIMSS researchers used data from multiple sources to reinvestigate course content and the intended curriculum. Comparing a random sample of U.S. state curriculum guides for grades 1 through 12 with curriculum guides in nearly 50 other countries, they identified the broad stretch of the U.S. science and mathematics intended curriculum.[14] The data confirmed the curriculum guides' lack of focus in U.S. math-

[14]TIMSS looked at the intended curriculum by analyzing the topics included in a random sample of 22 mathematics and 18 science standards documents from state and national curriculum associations. To compare the United States with other countries, the researchers created a composite score that represented the topics covered in the different guides reviewed. Creating this composite was necessary because the United States does not have a national curriculum. The composite therefore represents "the aggregate policies, intentions, and goals of the many educational subsystems making up the U.S. federation of guiding educational visions in the sciences and mathematics" (Schmidt, McKnight, and Raizen 1997, p. 13).

ematics and science curriculum: U.S. teachers in grades 1 through 12 intended to cover more topics than teachers in other countries. Figure 3.1 shows that in mathematics:

- The United States ranked above the 75th percentile among 36 countries in topic coverage for grades one through eight.
- The United States dropped below the 75th percentile for grades 9 through 12.
- Germany and Japan, by comparison, ranked at or below the median in at least eight grade levels and ranked near or below the 25th percentile in at least six grade levels.

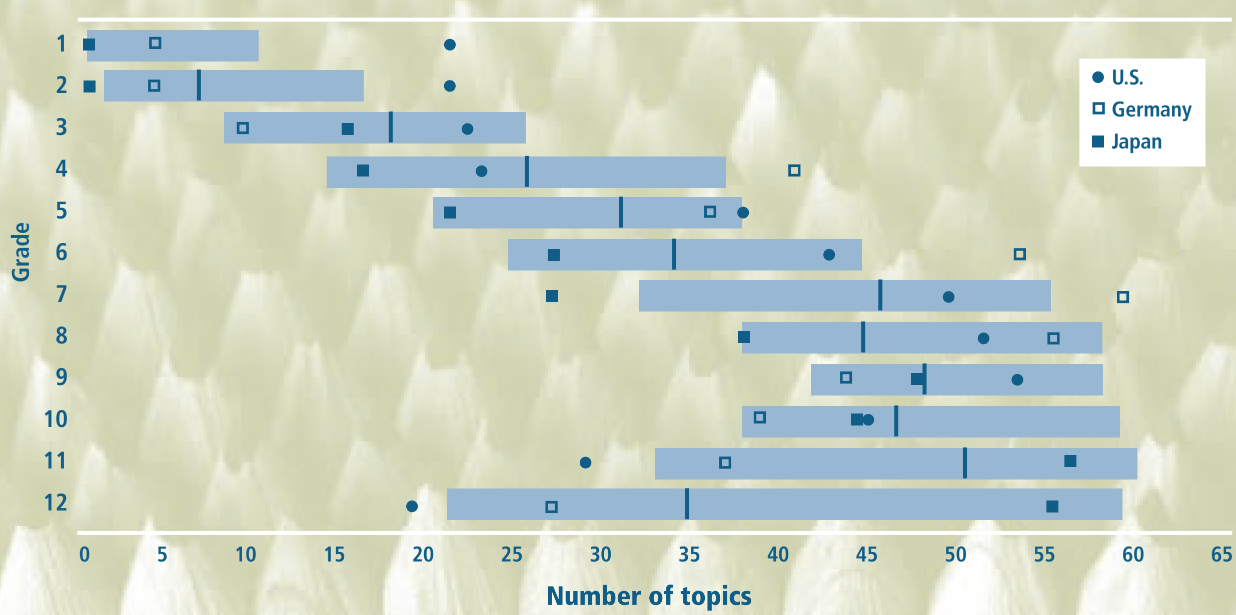
Figure 3.2 shows the following about science:

- The United States scored above the median in grades one through nine and above the 75th percentile in grades one and two among 35 countries.
- The United States scored at or below the median in three grade levels (10 through 12), while Germany scored at or below the median in six grade levels and Japan in 10 grade levels.

This is a critical finding. Standards documents serve as official education policy both in this country and abroad; they signal educators about what they are expected to do in their classrooms. These findings echoed the SIMS concerns from a decade earlier:

> These curricula [in the United States] express an intention to do something of everything and, on average, less of any one thing ... While the relationship between curricular focus and achievement may remain conjectural, the con-

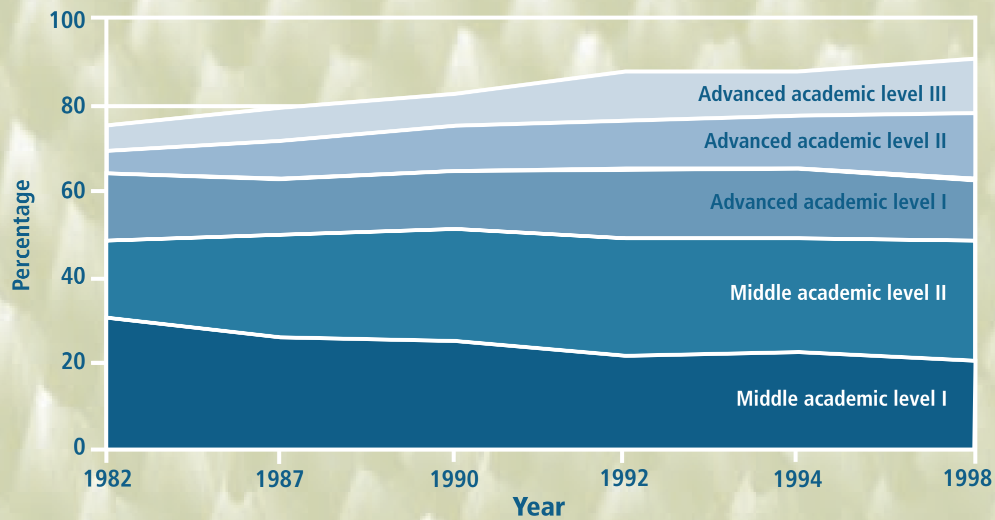
Figure 3.2 Number of science topics intended to be covered in 35 countries



NOTE: The light blue bars show how many science topics were intended to be covered at each grade in the TIMSS countries. The bars extend from the 25th percentile to the 75th percentile among countries. The blue line indicates the median number of topics at each grade.

SOURCE: William H. Schmidt, Curtis C. McKnight, and Senta A. Raizen, *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education,* Third International Mathematics and Science Study, 1992–93.

Figure 3.3 Percentage distributions of high school graduates according to the highest levels of advanced mathematics courses taken: 1982, 1987, 1990, 1992, 1994, and 1998



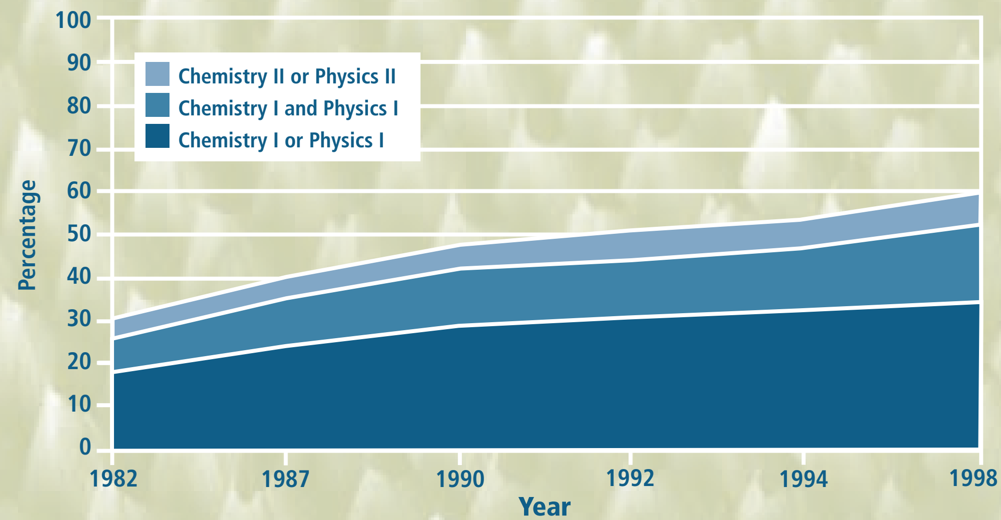
SOURCE: U.S. Department of Education, NCES. High School and Beyond Longitudinal Study of 1980 Sophomores, "Second Follow-up" (HS&B: 1980/1984); National Education Longitudinal Study of 1988 Eighth Graders, "High School Transcript Study" (NELS: 1992); and 1987, 1990, 1992, 1994, and 1998 National Assessment of Educational Progress High School Transcript Studies.

> ventional wisdom (e.g., about time on task) suggests that lower achievement results are likely and that the United States is likely to suffer in cross-national comparisons (Schmidt, McKnight, and Raizen 1997, p. 51).

Compounding the lack of focus in the intended U.S. curriculum, the implemented curriculum varies tremendously from classroom to classroom in this country. For example, SIMS researchers found that 13-year-olds in the United States received very uneven exposure to a range of curriculum topics relative to the other countries (McKnight et al. 1987). This might be explained by the fact that during the time of the SIMS data collection, more than 82 percent of the U.S. students in this age group were in schools offering two or more differently titled mathematics classes (Schmidt et al. 1999). In contrast, many of the countries that scored high on the SIMS mathematics tests (France, Hong Kong, and Japan) offered the same mathematics course to *all* students.

Student enrollment patterns show the percentage of high school graduates who studied each subject. Figures 3.3 and 3.4 reveal trends in the mathematics and science enrollment patterns of high school students. From 1982 to 1998, there was an increase in the percentage of students enrolling in higher-level mathematics and science courses (Wirt et al. 2000). High school graduates in 1998 were more likely than their 1982 counterparts to have taken the most advanced mathematics courses, such as trigonometry, precalculus, and calculus. For example, 5 and 6 percent of 1982 high school graduates took these types of advanced academic-level II and III courses, respectively (as defined in Wirt et al. 2000), and in 1998, the percentage of students taking these courses rose to 15 and 12, respectively. Furthermore, in the middle academic level, a larger percentage of high school graduates took level II courses (such as algebra II) as their highest-level mathematics course in 1998 than in 1982 (increasing from 18 to 28 percent), and a

Figure 3.4 Percentage distributions of high school graduates according to the highest levels of advanced science courses taken: 1982, 1987, 1990, 1992, 1994, and 1998



SOURCE: U.S. Department of Education, NCES. High School and Beyond Longitudinal Study of 1980 Sophomores, "Second Follow-up" (HS&B: 1980/1984); National Education Longitudinal Study of 1988 Eighth Graders, "High School Transcript Study" (NELS: 1992); and 1987, 1990, 1992, 1994, and 1998 National Assessment of Educational Progress High School Transcript Studies.

lower percentage took level I courses (such as algebra I and plane geometry) as their highest-level mathematics course (decreasing from 31 to 20 percent) (Wirt et al. 2000).

In science, the trend is similar. For example, 5 percent of high school graduates completed chemistry II or physics II in 1982, but 7 percent completed these courses in 1998. Furthermore, during the same period, the percentage of graduates who completed physics I increased from 18 to 34, and the percentage of graduates who completed chemistry I increased from 7 to 19 (Wirt et al. 2000).

Despite these encouraging signs, the experience is not reflected equally among racial/ethnic and income groups. In 1998, white and Asian/Pacific Islander high school graduates were usually more likely than black, Hispanic, and American Indian/Alaskan Native graduates to complete advanced academic-level mathematics and the highest-level science courses (Wirt et al. 2000). For example, while 45 percent of white and 56 percent of Asian/Pacific Islander high school graduates took at least one advanced mathematics course in 1998, 30 percent of blacks, 26 percent of Hispanics, and 27 percent of American Indian/Alaskan Natives took courses at this level (Wirt et al. 2000). Finally, students from low-income families were less likely to be enrolled in a college preparatory track through which they would be more likely to take such courses (U.S. Department of Education 1995).

All of the available data have limitations, however. The TIMSS approach of examining standards documents reveals the intended curriculum: what states, districts, and schools are *asking* teachers to cover in their classrooms, not what actually gets covered. NCEE standards that rely only on course titles, with no guidance on the topics that should be taught or mastered within those courses, do not directly address the SIMS/TIMSS concerns that topic coverage is not uniform across classrooms and that the U.S. curriculum lacks focus. The NCES data on enrollment patterns attempt to address the issue at a lower, more revealing level, but they are also limited because they depend heavily on course titles that may incorrectly assume identical course content,[15] and they are not applicable to primary grades.[16]

Recognizing these limitations to curriculum reform, policymakers began to call for more detailed education "standards" in the 1990s, and many state and local governments and national curriculum associations have responded. At around the same time, researchers began to call for curriculum indicators that focus more closely on elements within the curriculum (Murnane and Raizen 1988; Porter 1991; Shavelson et al. 1987; U.S. Department of Education 1988). In response, studies have been launched to ascertain how to develop more precise curriculum content indicators that measure the intellectual rigor and cognitive demand of the implemented curriculum (Burkam, Lee, and Smerdon 1997; Burstein et al. 1995; Porter et al. 1993). Similarly, NCES has recognized the need for data that are more specific. National surveys (such as the National Assessment of Educational Progress and the National Education Longitudinal Study of 1988) now ask teachers if they cover various topics. Neither of these surveys addresses content coverage to the same depth as SIMS or TIMSS, but they have started to collect information on the implemented curriculum that goes far beyond the surface views of intended curriculum provided by course credits and titles.

B. INDICATOR 6: PEDAGOGY

The implemented curriculum as represented by instructional delivery or pedagogy is an increasing focus of researchers' attention because it illuminates the black box between educational inputs and student outcomes (Mayer 1999b). This is a relatively new research area. When earlier research failed to identify a consistent relationship between resources and student achievement, researchers began to investigate variation in how those resources (teachers, classrooms, materials) were combined. Striving to understand why identical inputs result in varied student outcomes, this new line of research on pedagogy assumes that teachers have different effectiveness because of how they approach the curriculum, select instructional actions for themselves and their students, and employ their resources.

National data on instructional practices can provide information about how quality classrooms deliver student learning, the prevalence and effectiveness of competing pedagogical approaches, and trends in instructional approaches over time. Such information can be used to gauge the degree to which the various techniques recommended by state and local governments and national curriculum associations have been implemented and whether those techniques actually do improve student learning.

At present, there is little agreement about the relative effectiveness of two prevalent types of pedagogical approaches, and improved national instructional practices data

[15]Do all algebra I classes cover the same material for the same duration? In fact, findings from SIMS and TIMSS suggest they do not. And to what extent have course titles, but not substance, changed to better "meet" external standards?

[16]Course titles are virtually useless as indicators of the curriculum in the primary grades because students at this level do not have electives and the course titles are generally no more specific than "reading," "mathematics," "science," and "social studies."

might help inform this debate. Proponents of traditional instruction, such as drill and practice activities in which students work toward skill mastery, suggest that this approach provides teacher-directed control and structure of classroom activities. Summarizing the findings from numerous studies, Brophy and Good (1986) cite elements of this approach as "essential":

> Finally, it should be stressed that there are no shortcuts to successful attainment of higher-level learning objectives. Such success will not be achieved with relative ease through discovery learning by the student. Instead, it will require considerable instruction from the teachers, as well as thorough mastery of basic knowledge and skills which must be integrated and applied in the process of higher level performance. Development of basic knowledge and skills to the necessary levels of automatic and errorless performance will require a great deal of drill and practice. Thus, drill and practice activities should not be slighted as low level. They appear to be just as essential to complex and creative intellectual performance as they are to the performance of a virtuoso violinist. (p. 367)

Advocates of constructivism, on the other hand, assume that "learning occurs as students actively assimilate new information and experiences and construct their own meanings" (National Council of Teachers of Mathematics 1991, p. 2). Because teachers must help students construct their own meaning, passive activities such as lectures and seatwork are considered significantly inferior teaching approaches. According to this line of research, the most effective instructional approach is one that places less emphasis on memorization of facts and mastery of routine skills and greater focus on application, reasoning, and conceptual understanding. A pedagogic approach emphasizing group work and class discussion, for example, is considered essential because it engages the students while offering the teachers a view of the students' thought processes (Case and Bereiter 1984; Cobb and Steffe 1983; Davis 1984; Hiebert and Wearne 1993; Lampert 1986; Lesh and Landau 1983; Schoenfeld 1987).

There are very few national data on the relationship between pedagogy and learning that might illuminate this debate on which approach is more effective perhaps in part because the push for routine collection of classroom instructional process data began only in the late 1980s (see, for example, Murnane and Raizen 1988; Porter 1991; Shavelson et al. 1987; U.S. Department of Education 1988). In 1994–95, NCES collected instructional process data from a sample of K-12 teachers and published the data in 1997 in a new "instructional methods" category in its list of annually released vital statistics (Smith et al. 1997). NCES data show, for example, that at least once a week 98 percent of teachers provide whole group instruction, 86 percent of teachers work with small groups of students, 73 percent use manipulatives or models to demonstrate a concept, and 40 percent use portfolios to assess their students' work in English/language arts (Henke, Chen, and Goldman 1999). These data also provide a way to judge the extent to which various pedagogical approaches are used in classrooms and how these trends compare with the teaching approaches recommended by the research or by various curriculum standards documents.

In another NCES effort to understand pedagogy and to experiment with how best to collect such data, TIMSS researchers supplemented written teacher surveys with a video survey of eighth-grade mathematics teachers in action in the United States, Germany, and Japan (Stigler et al. 1999). The video survey, like classroom observations, promises specificity and objectivity (in that the data are not self-reported by teachers) and has the promise of being available for systematic scrutiny by a wider research audience.

While these approaches all show promise for collecting informative data on pedagogy, they each have limitations (Mayer 1999a). Some survey data have been found to be unreliable. Perhaps because the teaching process consists of complex interactions between students and teachers that a survey cannot accurately depict, teachers may provide biased responses to a survey because they feel that they should respond to the questions in an "acceptable" or "socially desirable" way, and teachers may unknowingly provide misleading responses to survey questions (Henke, Chen, and Goldman 1999; Mayer 1999a; Mayer 1999b). On the other hand, case-study data may be valid, but they are expensive and do not provide a portrait beyond the few classrooms studied (Burstein et al. 1995). And video data collection has validity threats, including observer effects, sampling concerns (determining the appropriate number of individual teacher observations and accounting for differences in content and practice at different times during the school year), and confidentiality questions (Stigler et al. 1999).

Despite these current limitations, there remains intense interest in better understanding how variation in teachers' methods of classroom instruction translates into variation in student learning. Recognizing the importance of the information yet the weakness of the data, researchers have begun to design surveys that will gather more precise information about instructional practices (Burstein et al. 1995; 1998; Mayer 1999b; Mullens and Gayler 1999). Continued collection and analysis of pedagogy data, especially experimental design data linked with student outcomes, will improve our knowledge of what pedagogy works, when it works best and why, and how best to measure it.

C. INDICATOR 7: TECHNOLOGY

Education Secretary Richard Riley's statement that "over the last decade, the use of technology in American life has exploded" (U.S. Department of Education 1996, p. 1) characterizes the inevitable: Computers and the Internet play a role in classrooms. And though the explosion has yet to reverberate within schools to the same degree it has in other sectors of American society, technology and computers are rapidly appearing in schools and redefining the perception of a quality school. Computers and Internet access are used in a variety of ways in schools, and each use may have an independent impact on student learning. Very little research on the effect of technology on learning looks at the uses and effect of Internet access; most examines the instructional power of the computer to teach discrete skills. Numerous studies conducted in the elementary and secondary grades have concluded that student learning is enhanced by computers when the computer is used to teach discrete skills in the style referred to as "drill and practice." Four separate meta-analyses support this conclusion, and so far none refutes it (President's Committee of Advisors on Science and Technology 1997). The magnitude of the effect is considerable, ranging in these four studies from 0.25 to 0.40 standard deviations. The benefits appeared to be strongest for students of lower socioeconomic status, low achievers, and those with certain learning problems (President's Committee of Advisors on Science and Technology 1997).

In the past decade, education reformers have seized on research about how students learn (see Indicator 6: Pedagogy, above) and on research about what skills are demanded in today's economy (for example, Murnane and Levy 1996b) to conclude that students need to be taught differently. Research on the application of computers for developing higher-order thinking skills problem-solving, group work, and hands-on learning activities, however, is less extensive and less conclusive. Two studies show positive effects (Glennan and Melmed 1996; Wenglinsky 1998), while a third concludes it is not known whether computers can be used for this type of teaching in a cost-effective manner with any "degree of certainty that would be desirable from a public policy viewpoint" (President's Committee of Advisors on Science and Technology 1997). It may be that these studies are less conclusive because teachers are less adept at teaching in this new way. This suggests that the challenge of using computers to develop higher-order thinking skills may be closely tied to the challenge of improving professional development for teachers.

Because computer technology is expensive, researchers and administrators want to assess its availability, equity, use, and effect on student learning. Recognizing both the growing use of technology in schools and the limited amount of applicable research and data collection, President Clinton's Panel on Educational Technology identified a "pressing need for large-scale, federally sponsored research and evaluation" on school technology (President's Committee of Advisors on Science and Technology 1997, p. 51).[17] Secretary Riley echoed that concern when speaking to the 1999 National Conference on Educational Technology.

Thus, large-scale data collection on technology in the classroom should address four issues: the availability of quality computers and Internet access, whether computers and access are equitably distributed among schools and students within schools, how the technology is used, and how it affects student learning.

To what extent are computers and Internet access available to students? On the surface at least, there is little doubt that computer technology is becoming increasingly available in schools. In 1999, there was an average of 6 students for each computer from a 125 to 1 ratio in 1983 (Coley, Cradler, and Engel 1997; U.S. Department of Education 2000b). If that pace of improvement continues, achieving the optimum 5 to 1 student-to-computer ratio suggested by the U.S. Department of Education (U.S. Department of Education 1996) seems likely. This overall average student-to-computer ratio, however, hides two facts: the distribution of computers per student is skewed, and many computers included in that count may be old and have limited usefulness. In 1994, for example, 4 percent of the nation's schools had at least one computer for every 4 students, while 46 percent of the schools had one computer per 16.5 students, and 10 percent of the schools had one computer per 28.5 students (Glennan and Melmed 1996). A 1998 national probability sample of 655 elementary and secondary schools revealed that "over half of the computers are out of date.... And in elementary schools almost two-thirds are of limited capacity" (Anderson and Ronnkvist 1999, p. 5).[18] This is potentially problematic because older computers do not have the capacity to link to the Internet or run cur-

[17]Almost all data inventorying the computer hardware in the nation's schools come from two private companies: Quality Education Data (QED) and Market Data Retrieval (MDR). QED data have been used in several major reports (see, for example, Coley, Cradler, and Engel 1997; Glennan and Melmed 1996; President's Committee of Advisors on Science and Technology 1997; U.S. Department of Education 1996). These companies survey schools to assess their equipment inventories and future buying plans. They attempt to get data from every school because their main objective is to create marketing lists and reports for technology companies. Because schools are under no obligation to complete the surveys, many schools choose not to, and therefore the inventory data present a conservative picture—a limitation that is generally overlooked due to heightened policy interest in these areas.

rent multimedia applications such as CD-ROM reference and encyclopedia programs. They can be used to perform drill and practice sessions and to develop keyboard skills.

The Internet is thought to provide a variety of rich learning opportunities, and therefore schools have been rapidly connecting to it. Internet access existed at 35 percent of public schools in 1994, but soared to 95 percent by 1999 (U.S. Department of Education 2000b). In 1999, in 37 percent of schools, however, access to the Internet existed at only one location thus making regular instructional use difficult in these schools (U.S. Department of Education 2000b).[19]

Is there an equitable distribution of computers and Internet access among schools with high and low concentrations of poverty? A recent study (Anderson and Ronnkvist 1999) found that schools with high concentrations of poor or minority students have fewer computers and are less likely to have Internet access. While nationally representative data suggest this gap is narrowing, they also show that "large gaps … in the quality of the computer equipment available" still exist (Anderson and Ronnkvist 1999, p. 16). The gap also exists regarding Internet access. For schools with more than 70 percent of students eligible for free or reduced-price lunch, 39 percent of all instructional rooms had Internet access compared with 62 to 74 percent for schools with lower concentrations of poverty (U.S. Department of Education 2000b).

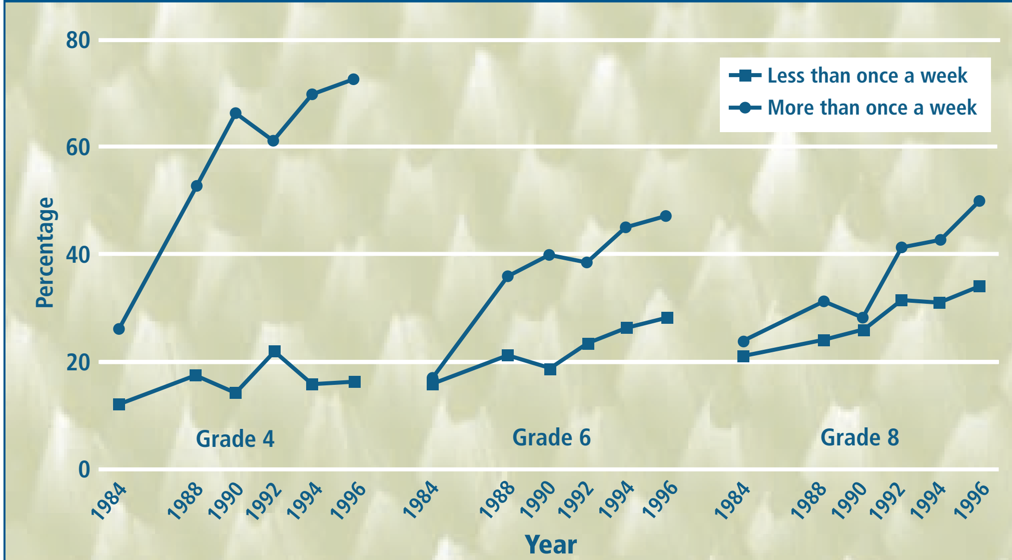
In addition to understanding the availability of technology in schools, researchers are interested in the frequency and type of computer use by students in school (Becker, Ravitz, and Wong 1999). Data from the 1998 National Survey on Teaching, Learning, and Computing indicate that many students are "searching for information and producing written and visual products that reflect their cognitive and creative effort" (Becker, Ravitz, and Wong 1999, p. 13). Word processing is the most common computer use across grade levels: 66 percent of elementary teachers and 45 percent of high school teachers report that students use computers for word processing in their classes. Two other uses of computers are prominent as well. Fifty-four percent of elementary teachers report that students use CD-ROM reference software, and 34 percent of high school teachers report that students use the World Wide Web in class.

Data from the National Assessment of Educational Progress depicted in Figure 3.5 suggest that computer use in schools has been on the rise (Smith et al. 1997). As shown in Figure 3.6, students report spending more time using computers to "learn things" and to "write stories" than to "play games" (U.S. Department of Education 1997). The data on computer use are limited for three reasons. First, there is only general information about the types of computer use and no information on the level of cognitive demand required by students when using a computer. Thus, it is impossible using any single source to distinguish whether students are playing games, learning to use the Internet, or learning the core curriculum through specialized software at school or home. Second, the data fail to identify the length of time students spend on computers, which makes it difficult to assess the impact of that technology on student learning. Finally, the data sometimes fail to distinguish between home and school usage, making it difficult to attribute changes to improvements in school quality.

[18]The authors suggest that elementary schools may have fewer modern machines because middle and high schools tend to "pass on their replaced computers to the elementary schools" (p. 5).

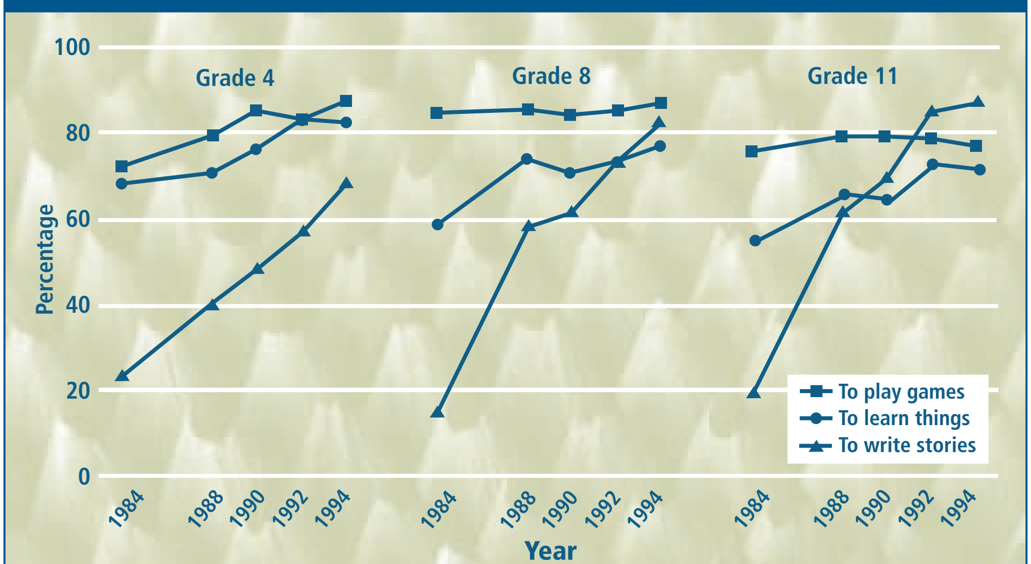
[19]Internet access in private schools lags. About one-quarter of private schools and 5 percent of their instructional rooms had Internet access in 1995, increasing to 67 percent and 25 percent, respectively, by 1998 (U.S. Department of Education 2000a).

Figure 3.5 Percentage of students who reported using a computer at school and frequency of use, by grade: 1984–1986



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, *Almanac: Writing, 1984 to 1996*, 1998. U.S. Department of Commerce, Bureau of the Census, October Current Population Surveys.

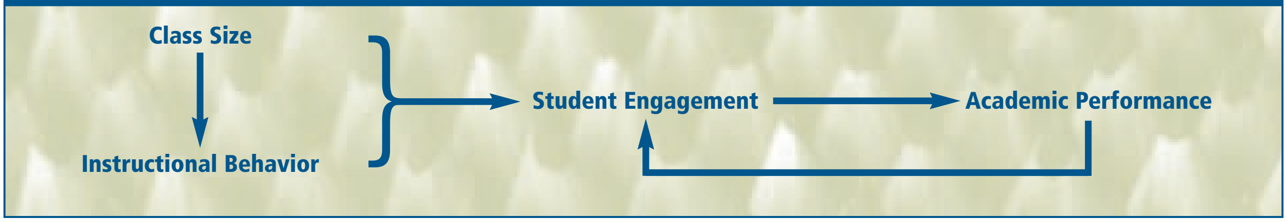
Figure 3.6 Percentage of students who reported using a computer to play games, learn things, and write stories, by grade: 1984–1994



SOURCE: U.S. Department of Education, National Center for Education Statistics, *The Condition of Education 1997*, 1997. U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, *Almanac: Writing, 1984 to 1994*, 1996.

Currently, the available national information on classroom technology is skewed: We have ample data on the availability of computers and some information about their distribution among schools and students, but we know little about how computers or the Internet are used. The country would benefit from more and better information on the

Figure 3.7 Relationship between class size and academic performance



NOTE: This is a modification of a figure presented by the U.S. Department of Education (1998) in "Class Size and Students at Risk: What is Known?" U.S. Department of Education, Office of Educational Research and Improvement, National Institute on the Education of At-Risk Students.

instructional role that computers play in classrooms and their effectiveness in achieving student learning.[20]

D. INDICATOR 8: CLASS SIZE

How teachers implement the course content, instructional pedagogy, and technology use in the classroom may all be influenced by the number of students in the class. Manipulating class size as a way to improve student learning is now at the forefront of the education policy debate because some evidence suggests that students may achieve more in smaller classes, particularly primary-grade students who are minorities or who come from economically disadvantaged backgrounds (Grissmer, Flanagan, and Williamson 1998; Krueger 1998; Mosteller 1995; Mosteller, Light, and Sachs 1996; U.S. Department of Education 1998). Why does class size matter, and how might it affect student learning? With fewer students, teachers may be able to employ a different pedagogical approach and implement the curriculum in a more effective manner. Figure 3.7 presents a model proposed by the Office of Educational Research and Improvement for how class size may affect student achievement. In the model, class size influences instructional behavior, and both class size and instructional behavior have an impact on student engagement. Higher levels of student engagement lead to higher achievement and vice versa, creating a cycle that leads to greater learning and productivity (Finn 1998).

The number of studies examining the relationship between class size and teaching quality are limited, and a better understanding is needed of the effects of class size on instructional practice and the quality of the classroom environment. From the limited research available, there is evidence that teachers in smaller classrooms deal with fewer disciplinary problems, spend more time on instruction and enrichment activities, and offer more opportunities for student participation (Achilles 1996; Finn 1998; Shapson et al. 1980; U.S. Department of Education 1998). Teachers also express greater satisfaction and report better morale when working with smaller classes (Shapson et al. 1980; U.S. Department of Education 1998).

Do smaller classes also lead to greater learning? Hundreds of studies focus on the relationship between class size and student achievement, and several researchers have compiled and analyzed these studies. Most of these meta-analyses have concluded that smaller classes do raise student test scores. Glass and Smith (1979) concluded that stu-

[20]It should be noted, however, that new data on the instructional role of computers is being made available in an NCES report (Smerdon and Cronen 2000) that was released as this report went to press.

dents in classes of 20 or fewer students perform better academically than students in larger classes. Robinson and Wittebols (1986) sharpened the focus by contending that small classes show stronger benefits for disadvantaged and minority students in kindergarten (Finn 1998; Robinson and Wittebols 1986; U.S. Department of Education 1998). Slavin (1989) used a different set of studies that looked at class-size reductions over a year or more and that compared regular-sized classes with those that were at least 30 percent smaller and had 20 students or less and also found that small classes increased student learning. He estimated the median effect size of attending a small class rather than a large class has a 0.13 standard deviation increase in student test scores; this effect did not seem to follow students when they returned to regular-sized classrooms (Finn 1998; Slavin 1989; U.S. Department of Education 1998).

Others disagree. Their views are represented in this summary statement by Hanushek:

> Extensive statistical investigation of the relationship between class size and student performance shows as many positive as negative estimates. With close to 300 separate estimates of the effect of class size, there is no reason to expect performance improvements from lowering class sizes. (Hanushek 1998, p. iii)

However, in a reanalysis of Hanushek's (1989) synthesis of education production function studies, Hedges et al. (1994) concluded that class size made a difference in student learning. In addition, in a subsequent study that examined a larger and, according to the authors, more comprehensive group of studies, Greenwald et al. (1996) again found a positive, statistically significant relationship between small classes and student performance.

These findings are supported by a recent and widely cited study on class size, a randomized experiment conducted by the state of Tennessee (Krueger 1998; Mosteller 1995; Sanders and Rivers 1996). Because randomized studies are considered to be among the most valid study designs and because they are so rare in the field of education research, the results from this study have received a tremendous amount of attention from researchers, policymakers, and the general public.[21] The first phase of the study, called Project STAR (Student-Teacher Achievement Ratio), compared the achievement of elementary students in small classes with those in regular-sized classes with and without a full-time teacher's aide.[22] By the end of first grade after two years of intervention, students attending smaller classes benefited by approximately 0.25 of a standard deviation. The effect sizes for the Stanford Achievement Test were 0.23 for reading and 0.27 for math.[23] This means that an average student achieving at the 50th percentile rank would see a 10 percentage point boost in achievement (Mosteller 1995). These effect sizes show the advantage of small classes over regular-sized classes in standard deviation units.

[21]In an experimental study, the researchers are able to allocate each student to a control or treatment group, typically through random assignment whereby each student has an equal chance of being in either group. This produces two statistically equivalent groups, allowing researchers to attribute any differences in outcomes to the intervention. In an observational study, the researchers have no control over how people choose from among various options, a phenomenon called self-selection. When self-selection occurs, it is not possible to determine whether groups differ in outcomes because of the intervention or because of other characteristics that come about as a result of self-selection. Experimental studies are therefore thought to be superior to observational studies for understanding causation.

[22]Over four years, about 80 schools and 11,600 students participated in Project STAR. A small class had 13 to 17 students per teacher and a regular class had 22 to 26 students. Students and teachers were assigned to each type of class at random, and researchers made sure that a school with at least one small class also had a regular-sized class with or without an aide in an attempt to rule out the effects of school characteristics when analyzing results.

When analyzed by race/ethnicity, the results showed that black students experienced a greater boost in achievement than white students in the early years. The effect size for black students was double the effect size for white students when standardized achievement test results were averaged over the first two years (Mosteller 1995). Further examination of the Tennessee data confirmed the finding that black and disadvantaged students benefited disproportionately in the early grades (Achilles 1996; Finn 1998; Grissmer, Flanagan, and Williamson 1998; Hanushek 1998; Krueger 1998; Mosteller 1995; Mosteller, Light, and Sachs 1996; U.S. Department of Education 1998). This finding also coincides with a recent analysis of data from the National Assessment of Educational Progress that shows the trend toward smaller classes may have contributed to the decrease in the achievement gap between nine-year-old blacks and whites (Grissmer, Flanagan, and Williamson 1998). However, in which grade Tennessee students experienced the greatest gains and how long those gains have persisted continue to be debated (Achilles 1996; Hanushek 1998; Krueger 1998; Mosteller 1995).

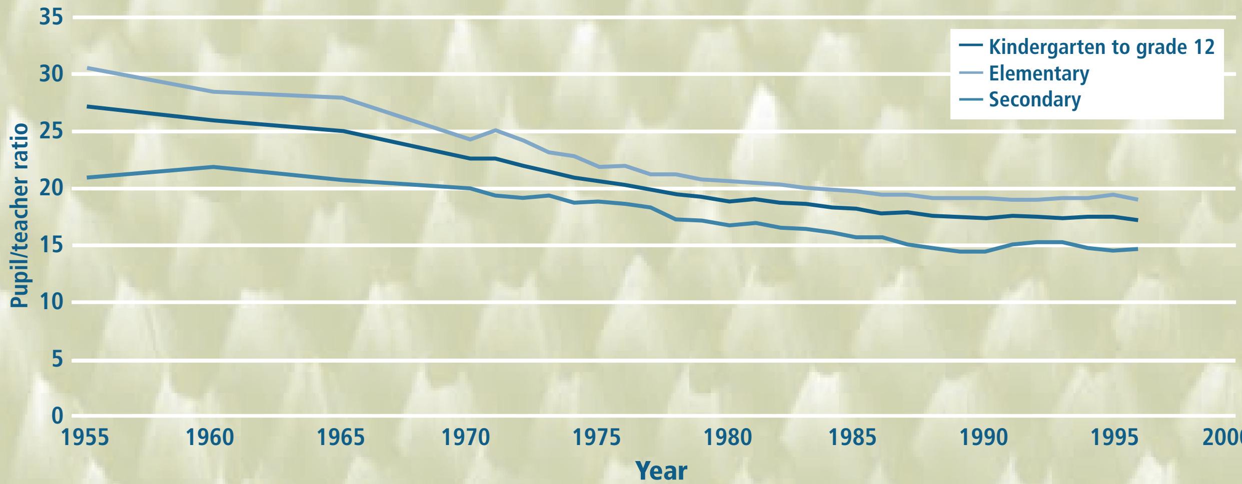
What is the optimal class size in the primary grades, or the threshold at which gains in student achievement are realized? Glass and Smith (1979) and Slavin (1989) have suggested that the threshold lies somewhere between 15 and 20 students; the Tennessee study compared classes of 13 to 17 students with classes of 22 to 26 students. How do these class size numbers compare with classrooms across the country? In 1994, more than 47 percent of elementary classes in the United States had 25 or more students (*Education Week* 1998), and in 1998 the average public elementary school class consisted of 23 students (Lewis et al. 1999). Thus, the majority of elementary classrooms is far from being in the 13- to 20-student range that these researchers identify as desirable.

Yet, while these studies suggest there may be a link between class size and student achievement, such as depicted in Figure 3.7, less is known about how that link might function and under what conditions. Murnane and Levy (1996a) found that reducing class size in Texas classrooms made a difference in student learning only where there was a school or district effort to change teaching practices and where teachers actually changed their instructional methods. And even when class size reduction is shown to be effective in raising student achievement, the additional teachers required to reduce class size will be costly (Finn 1998; Odden 1990; U.S. Department of Education 1998).

Furthermore, large-scale efforts to reduce class size (with or without professional development on appropriate instructional techniques for smaller groups) may also set in motion unintended negative consequences. For example, two years after California's initiative for smaller classes in kindergarten through third grade greatly increased the demand for a limited supply of qualified teachers, more than 12 percent of California K-3 classroom teachers do not meet minimum qualifications compared with less than 1 percent before the initiative (Bohrnstedt and Stecher 1999). In addition, those underqualified teachers are more likely to teach minority and disadvantaged students. In fact, third-grade students in the bottom quartile of reading achievement in California are five times as likely to have an underqualified teacher as students in the top quartile (Shields et al. 1999). These studies suggest that lowering class size by itself may not

[23]Effect sizes were calculated by subtracting the average achievement score for regular classes from the mean score for small classes in standard deviation units. The average achievement score for regular classes was obtained by taking the average of the mean score for regular classes with aides and without aides (Finn 1998). The Stanford Achievement Test is a standardized achievement test that was administered to all students in the study at the end of each school year

Figure 3.8 Public elementary and secondary pupil/teacher ratios, by grade level: Fall 1995 to fall 1996



SOURCE: U.S. Department of Education, National Center for Education Statistics, *Statistics of Public Elementary and Secondary Day Schools;* Common Core of Data Surveys; and *Projections of Education Statistics to 2008.*

enhance student learning if instructional practices remain unchanged or if unqualified teachers are used to reduce class size.

What do we know about national estimates of class size? Nationally representative data show that the pupil-teacher ratio across all public elementary and secondary schools was 27 pupils to 1 teacher in 1955 and 17 to 1 in 1991 (see Figure 3.8). Pupil-teacher ratios, however, are an imprecise indicator of class size because they do not account for frequently intense teaching resources targeted toward Title I, special education, or bilingual services in some schools. When schools use those resources to attain low ratios for those targeted populations, they mask much higher ratios for the remaining classes. This is especially true for poor and urban districts because they enroll higher numbers of students needing specialized instruction (Finn 1998; Hanushek 1998).

The better indicator of class size is the average number of students in regular classrooms. Unfortunately, national class size estimates (distinct from pupil-teacher ratios) date back only to the late 1980s. The average class size in public elementary and secondary schools was similar in the 1987–88, 1990–91, and 1993–94 school years (Smith et al. 1997). In 1993–94, the average public elementary class consisted of 24.1 students, while the average secondary class consisted of 23.6 students. As of 1998, the average class size for public school general elementary classrooms was down to 23 (Lewis et al. 1999).[24] Class-size data for secondary schools in 1998 are not available.

These data and the studies describing the effects and potential consequences of reducing class size highlight the need for more data and for analyses that examine the effects of class size on instructional practice and student learning that can result in more pre-

[24]General elementary classrooms are classrooms that are self-contained for most of the day and in which one teacher is responsible for instructing in all core subjects, including reading/language arts, mathematics, social studies, and science (but not necessarily art, music, and physical education). General elementary classrooms are distinct from departmentalized settings in which a teacher's main teaching assignment is in one particular subject area.

cise estimates of optimal size. Barriers to gaining a complete understanding of how class size is related to student achievement are a lack of consistent and appropriate data on class size, a lack of agreement on what methodology works best in analyzing the existing literature, and a lack of resources needed to conduct well-designed, random experiments. Despite these barriers, the preponderance of research suggests that reductions in class size have the potential to help all students in the primary grades.

E. SUMMARY

Influential studies over the past two decades cite a lack of focus in the intended school curriculum as a problem that may be related to the poor academic performance of U.S. students relative to students in other nations. Although a review of high school course titles between 1982 and 1994 suggests that this problem might be abating, reliance on course titles only touches the surface of the intended curriculum.

Nationally representative data that describe the implemented curriculum and the process of schooling are just becoming available. These data will help researchers study the instructional practices of teachers, the implemented curriculum, and the role that computers play in delivering the curriculum. As these data become available, new insights might arise concerning the role of curriculum in determining school quality. For example, these data might help determine whether differences in curriculum coverage help explain why small classes appear to benefit children in the primary grades and could help determine exactly what teachers do differently in small classes.

IV. SCHOOLS

This chapter builds on the previous discussions of teachers and classrooms by identifying school-level characteristics that contribute to a school's overall effectiveness. Education researchers, teachers, administrators, and parents all know that schools distinctly matter in the educational lives of children, that school effectiveness varies from institution to institution, and that much of what we know about the effectiveness of a school derives from an aggregate understanding of how that school fares on the important teacher and classroom indicators discussed in the preceding chapters. Beyond that, however, researchers have also begun to develop a literature on school-level characteristics that appear to play key roles in adding value to a school's quality.

How schools approach these aspects establishes the context for effective instruction and quality learning and may distinguish truly high-quality schools. Given this, school reform advocates and researchers have been striving to identify and replicate effective school-level approaches that lead to a quality education for all students (Bryk, Lee, and Holland 1993; Hanushek, Kain, and Rivkin 1999; Lee and Smith 1996; Marks, Doane, and Secada 1996; Phillips 1997). This school reform experimentation and research directed toward understanding the definitive attributes of effective schools blends new ideas about school organization and teaching with an improved understanding of educational processes. Results have been applied in comprehensive research-based school improvement models such as High Schools That Work, the Coalition of Essential Schools, and Accelerated Schools. Such research and experimental applications have sustained and focused the discussion on promising aspects of effective schools.

We used a two-step process to identify the most critical school-level characteristics. First, we researched the literature to identify the characteristics of schools that recent and current studies suggest are related to effective schools. Second, we examined how those attributes are measured or might be measured to assess their relationship with student learning. To be included in this list, characteristics had to occur at the school level, have a plausible relationship to student learning, and, where possible, be measured at the school level. Five school characteristics meet these criteria:

- School leadership that provides direction, guidance, and support,
- School goals that are clearly identified, communicated, and enacted,
- A school faculty that collectively takes responsibility for student learning,
- School discipline that establishes an orderly atmosphere conducive to learning, and
- School academic organization and climate that challenges and supports students toward high achievement.

In discussing these indicators, the chapter examines how to identify, measure, and understand the priorities that schools place within and between these five characteristics as they strive to best educate their students.

This discussion has structural and data limitations. First, quality schools are not simply a collection of discrete attributes that can be cut and pasted into operation. Instead, researchers and educators suggest quality schools are well-managed, amalgamated systems in which each action or policy directly links with and integrally affects all other parts (Bryk, Lee, and Holland 1993; Chubb and Moe 1990). Delineating the school-level influences into discrete chunks to determine the individual contribution of each to student learning may not adequately reflect that interconnected reality. For the sake of clarity in this analysis, however, we deconstruct the parts and examine them separately. Indeed, most school effects studies examining variation in student learning associated with structural or organizational conditions follow the same process (Bryk, Lee, and Holland 1993; Chubb and Moe 1990). They simultaneously acknowledge the integrity of the school as an integrated system and look within that system at individual school characteristics. While this is an effective analytic method for examining complex relationships, we must not forget the very real connectedness of the parts.

Second, as described in the preceding chapters, student learning is thought to occur primarily as a result of students' interaction with teachers, other students, and the curriculum. It may not be a direct or primary result of any school-level attribute or system. Bryk et al. (1993) agree with this by theorizing that most attributes of a school (measured at that level) are related only indirectly to student learning. School-level systems may have an indirect effect on student learning by establishing high expectations for educational experiences and by setting the context within which quality interactions can occur.

Third, even if certain school attributes indirectly influence student learning, identifying how they operate and collecting data to measure their statistical relationships with student learning is a complex task that may inhibit our understanding of their effects. Some teacher and classroom attributes identified in the previous chapters, such as teacher academic ability and class size, are easier to measure and quantify. Quantifying human incentives, motivations, opinions, and actions to estimate the effects of school-level attributes such as leadership, professional community, and organizational philosophies is not yet, and may never be, an exact science. So the validity of the findings is heavily dependent on measurement, specification, and interpretation issues that might easily cloud the results and inhibit rather than extend our understanding (Chubb and Moe 1990; Willms and Echols 1992). This is why most of the research on school-level attributes has been through case-study research and why very little analysis of national data reliably relates school-level characteristics to learning outcomes.

Finally, as will become apparent by the paucity of figures in this chapter, our understanding of school quality is hampered by a lack of school-representative and reliable national data on these five attributes. The data that are available may suggest the general status of an indicator, but sometimes they lack the reliability and specificity necessary to understand confidently what really matters about the indicator. Available data on school goals, for example, provide information on the relative importance of seven general goal statements, but no information on the issues that, research suggests, differentiate effective schools: how goals are identified, communicated, implemented, and institutionalized (Purkey and Smith 1983). Some of the available data on school discipline incidents and academic environment are an exception.

Despite these analytic and measurement challenges, school-level characteristics are thought to be important elements of quality schools. For each of the five school-level characteristics identified, this chapter describes the relevant research indicating why that characteristic is thought to be important, suggests how that indicator might be measured, presents, where available, nationally representative data that describe the current state or prevalence of that school characteristic, and identifies, where necessary, the type of nationally representative data necessary to better identify quality schools.

A. INDICATOR 9: SCHOOL LEADERSHIP

Schools need a competent individual or group of individuals who can provide direction, guidance, and support in the school's journey toward achieving its goals (Newmann and Wehlage 1995). Leadership provides a unifying focus, the impetus to work toward school goals, and a locus for decision making along the way. In quality schools, an individual or group of individuals takes responsibility to provide school leadership, assemble a faculty with the skills to achieve school goals, provide direct support for those teachers, and make teaching and learning a main preoccupation around which everything else revolves. A school without a leader is a collection of independent classrooms with individual goals and unconnected beliefs about what is important and how to achieve it. Pervasive and sustained student learning is more likely to occur in schools with leadership (Deal and Peterson 1998; Fullan 1998; Levine and Lezotte 1990). Some researchers, however, question whether school leaders make a measurable difference in fostering school quality. Hanushek et al. (1998), for example, suggest that constraints on school leaders, especially central office regulations on school organization and hiring in public schools, virtually eliminate any opportunity for school leaders to leverage change and improve school quality.

There is scant meaningful national data that illuminate these aspects of school leadership and support or refute either view. Most of our understanding about school leadership comes from case studies that can only suggest what might be true in a broader application. Those studies suggest critical issues of leadership: where it is located, how it is exercised, how it differs between public and private schools, and its relationship to student learning.

Theoreticians have used case studies to investigate where school leadership is located. Recent work has used a "power continuum" (King et al. 1996). At one extreme, power is consolidated in the principal, district personnel, or a small group of teachers. At the other extreme, decision making is shared and participants have equal access and voice. In the middle are schools where power is dispersed in small groups and communication is neither highly valued nor implemented (balkanized schools) and schools where teachers have high levels of individual autonomy but act on disparate goals (laissez-faire schools). If nationally representative data were available, classifying schools into these categories (or something similar) and relating the categories to student learning might meaningfully distinguish among leadership styles and highlight variation in student learning related to those styles.[25]

In addition to identifying where leadership is located, researchers suggest it is important to determine how it is exercised. What do school leaders actually do, and how are those actions related to student learning? Is it possible to identify those leadership actions that are consistently related to higher levels of student achievement? Two recent studies investigated behaviors common to principals of schools with high student achievement. Controlling for relevant principal and school background characteristics such as race/ethnicity, gender, and socioeconomic status, Louis et al. (1996) found that leaders in

[25]Does leadership style affect the ability of a school to reach its goals? King et al. (1996) say their data cited above suggest that schools were more likely to support high-level learning when the power was shared among the participants. The "broad participation, reciprocity, and collective focus on important issues characteristic of shared power" (p. 255) best facilitated reaching the goals of improved teaching and learning within the schools in their study.

schools with high student achievement "worked effectively to stimulate professional discussion and to create the networks of conversation that tied faculty together around common issues of instruction and teaching" (p. 194). In that process, they say, principals "delegated authority, developed collaborative decision-making processes, and stepped back from being the central problem solver" (p. 193) regardless of the formal definition of their role. These findings are congruent with results from other research (Leithwood, Jantzi, and Fernandez 1995; Louis, Kruse, and Associates 1995; Murphy 1994).

Work by Newmann and Wehlage (1995) took a similar approach, summarizing the characteristics of leaders in schools with high-achieving students. Such leaders, they say:

- Gave "central attention to building a schoolwide, collective focus on student learning of high intellectual quality,"
- Placed "issues of teaching and learning at the center of dialogue among the entire school community,"
- Gave concrete expression to "the norms and values" that comprise a school's vision,
- Created "time for reflective inquiry and opportunities for substantive staff development,"
- Saw themselves "at the center rather than at the top of their schools' organization,"
- "Shared power with staff and often with parents," and
- "Applied important political … and entrepreneurial … skills to relationships beyond the school." (pp. 291–292)

Is leadership style and process different in public and private schools? Available national teacher opinion data on leadership suggest striking differences. "Public school teachers were consistently less likely than their peers in private schools to agree" that the principals in their schools took the following actions (Henke et al. 1997, p. 82):

- Communicated expectations to staff,
- Was supportive and encouraging,
- Recognized staff for a job well done, and
- Talked with them about instructional practices.

Other national data suggest that such differences between leadership in public and private schools are likely to relate to structural variations in public and private school governance systems. Without the school district administration common to public systems, according to these analyses, private school principals are likely to shoulder more responsibilities (Bryk, Lee, and Holland 1993),[26] while in a public school the bureaucratic hierarchy dilutes the principal's leadership role (Chubb and Moe 1990).[27]

Information on the locus and exercise of school leadership that moves beyond opinion data would better position the education community to detect and understand mean-

[26]Catholic school principals act as the chief administrative officer in a position equivalent to an administrative combination of a public school principal and superintendent with responsibility for financial management, development and fund-raising, faculty selection and supervision, public and alumni relations, discipline, and instruction. They manage operations and personnel, build personal relations, and provide spiritual leadership (Bryk, Lee, and Holland 1993).

[27]Public school principals are "systematically denied much of what it takes to lead" and function more as lower-level managers than as leaders. The structure of the job within the public hierarchy promotes bureaucratic behavior and playing by the rules and may be especially enticing to individuals inclined toward those traits (Chubb and Moe 1990, p. 56).

ingful differences in leadership styles between principals in high- and low-quality schools and how those differences relate to variation in student learning. To understand the place of school leadership in developing quality schools, the education community would benefit from information on the placement of leadership (is it always in the school and is it always synonymous with the school principal?), how leadership is exercised (what do school leaders actually do in the process of leading?), and how specific leadership styles relate to student learning.

B. INDICATOR 10: GOALS

Whether called shared vision, shared beliefs, shared values, or common goals, a clear sense of purpose with participant buy-in is a key ingredient in any successful social organization, including schools (Purkey and Smith 1983). The process is neither simple nor clear-cut, however. Identifying school goals and achieving consensus among school administration, parents, staff, and community and then implementing and institutionalizing those goals can be an involved and difficult process. While the process may vary, researchers agree that successful schools begin by clearly identifying and communicating ambitious goals, then implementing and institutionalizing those goals with broad consensus from essential stakeholders (Bryk, Lee, and Holland 1993; Chubb and Moe 1990; Newmann and Wehlage 1995).

Researchers suggest that every step of the goals process is critical. After identifying appropriate goals, how those goals are articulated, perceived, and internalized by the school community may be independent factors. Whatever their intent or philosophical underpinning, the goals most likely to be achieved are those that are integral to every school activity (planning, teaching, assessing, decision making) and directly represented in the daily functioning of the school, from the mundane to the essential (graduation requirements, scheduling, advising, dress codes, faculty meeting agendas) (Bryk, Lee, and Holland 1993; Chubb and Moe 1990). Successful schools are those in which goals become part of the school fabric and all activities are aimed at achieving them. To be successful in practice, any structural features in place, such as block scheduling, interdisciplinary time, shared academic experiences, and projects that extend over several days or weeks, must be implemented to directly promote or support the school goals (Marks, Doane, and Secada 1996).

What do nationally representative data tell us about these questions about school goals? We know how public and private school principals responded to a list of seven general goal statements and how their responses have changed between 1990 and 1993 (Henke et al. 1996).[28] This information may be appropriate for knowing what the goals are and for tracking broad changes over time, but it lacks specific details on the issues that researchers suggest would help distinguish among effective schools. For that purpose,

[28]A majority of public school principals identified basic literacy, academic excellence, personal growth, and good work habits/self-discipline as a first, second, or third most-important school goal by more than twice as many as selected human relations skills, occupational or vocational skills, and specific moral values. Compared with principals polled in 1990–91, principals polled in 1993–94 were more likely (Henke et al. 1996) to indicate that academic excellence, occupational/vocational skills, and human relations skills were important goals and less likely (Henke et al. 1996) to indicate personal growth as an important goal (Henke et al. 1996). Specific moral values are an important school goal to a higher proportion of private school principals than public school principals (Fiore, Curtin, and Hammer 1997).

we need to know how school goals are identified, communicated, implemented, and institutionalized. The magnitude of this undertaking should not be underestimated. To understand the relationship between school goals and school quality, we would need rich contextual data from a nationally representative sample of schools on a wider range of more specific goals, on whether the school community agrees on and supports the goals, on whether the goals are implemented and represented in the daily functioning of the school, on how the goals change over time and in response to personnel shifts, and whether they relate to student achievement. This additional information would make it possible to assess schools' status relative to goals and perhaps to predict their success in establishing a positive context within which student learning can occur. Collecting this type of information, however, would be complex and costly.

C. INDICATOR 11: PROFESSIONAL COMMUNITY

Adults working collectively within a school can have a positive effect on student learning and school quality beyond their central and direct contribution as individual teachers. Research suggests that quality schools are likely to have a stable, professional community of experienced teachers with shared norms, values, and goals; a common focus on student learning; a willingness to collaborate; and an openness to reflection and new ideas, all directed toward high student achievement (Louis, Kruse, and Marks 1996). This suggests that teachers working together as a community of adults with individual and joint commitments to a set of common goals within the broader context of the school can have a powerful effect beyond their individual contributions (Louis, Kruse, and Marks 1996) and that this community effect may be related to improved student learning (Lee and Smith 1996). Students are the primary beneficiaries when a school community is cohesively focused and vigorously working toward common goals, but teachers also benefit. Teachers derive personal energy, fulfill their professional responsibilities, sustain quality performance, refresh their energy to be fully engaged with students, and stay focused on the central tasks of education. As the professional community works together to establish common goals and accomplish meaningful change toward those goals, the whole school is enriched (Louis, Kruse, and Marks 1996).

Additional research suggests that schools with an experienced and stable faculty may be more able to establish a professional community (Hanushek, Kain, and Rivkin 1998). High teacher turnover may negate the instructional continuity necessary for long-term learning gains, especially in mathematics. High School and Beyond data show that in schools with a high incidence of staffing problems, "students' behaviors are less oriented toward academics and sophomore achievement is lower" (Bryk, Lee, and Holland 1993, p. 223). Specifically, in schools with low faculty turnover, students "exhibit higher achievement in all curricular areas at both sophomore year and senior year." In schools where principals report higher staff absenteeism and lack of staff commitment and motivation, at-risk students may also face a greater likelihood of dropping out (Bryk and Thum 1989). None of the research determined, however, whether high faculty turnover is a cause of low-student achievement, an indicator of other more serious school problems that negatively affect both student learning and faculty retention, or both.

Is there less professional community at high-poverty schools? National data confirm that high-poverty schools have more inexperienced teachers, less-qualified teachers, and higher turnover. Rollefson and Broughman (1995 as cited in Henke, Chen, and Geis

2000) say that first-time teachers are overrepresented in high-minority schools.[29] Other information provided by 34 states indicates that while 38 percent of all teachers work in high-poverty settings, 64 percent of all unqualified teachers[30] work in high-poverty settings (U.S. Department of Education 1999b). Teacher turnover data from the Schools and Staffing Survey and the Teacher Followup Survey confirm these differences: High-poverty schools reported more difficulty hiring qualified teachers, were more likely to use long-term substitutes, were more likely to use pay incentives to obtain qualified teachers, and, when those efforts failed, were more likely to hire less than fully qualified teachers (Henke et al. 1997, p. 111).

While there currently exists reasonable national data on experience levels, turnover, and qualifications, additional school-representative national teacher data could provide information about other aspects of professional community: faculty agreement around school goals, the frequency and meaningfulness of teacher-to-teacher dialogue and collaboration, and faculty efforts to improve the teaching and learning conditions in the school. If these are related to student achievement as the research suggests, then analyses of expanded national data sets may be able to confirm expected links with student learning.

D. INDICATOR 12: DISCIPLINE

Researcher have found that a positive disciplinary climate is directly linked to high achievement (Barton, Coley, and Wenglinsky 1998; Bryk, Lee, and Holland 1993; Chubb and Moe 1990). An orderly school atmosphere conducive to learning could be an example of a "necessary, but not sufficient" characteristic of quality schools. Quality schools with high levels of student learning may have an accompanying high level of orderliness and discipline throughout the school as students are actively engaged in educationally productive activities. And yet an orderly school by itself is neither directly responsible for the student learning that might occur nor sufficient to ensure that learning will happen. Combined with other strong attributes of schools, teachers, and classrooms, however, a disciplined climate may be a necessary precondition that permits and perhaps enables good teaching and learning (Purkey and Smith 1983).

The issues that school discipline policies are designed to address are well known and range from the disconcerting to the dangerous. They include student disrespect for teachers, absenteeism, tardiness, use of alcohol and controlled substances, fighting, and possession of firearms. Less understood, however, is how discipline policy affects student behavior and how the resulting student behavior affects school quality and student learning.

Are some school policies more effective than others at controlling negative behaviors? Little information is available. Barton, Coley, and Weglinsky (1998), using the National Assessment of Educational Progress and the National Education Longitudinal Study data, suggest that the most effective policies to reduce the incidence of offenses in a school vary according to the targeted behavior. To reduce serious incidents including drug offenses, only a policy of severe punishment seems to be effective. "The more severe the punishment the lower the prevalence of the offenses" (p. 3).

[29]Nationally, one-fifth of all teachers work in high-minority schools, but one-third of all new teachers got their first job in a high-minority school (Henke, Chen, and Geis 2000).

[30]"Unqualified teachers" means teachers who do not meet state certification and licensure requirements and who have received an exemption from those requirements.

Whether the link between discipline and student learning is direct or indirect, it has an effect on school quality and student learning. Evidence comes from High School and Beyond and the National Educational Longitudinal Study. Bryk, Lee, and Holland (1993) indicate that a positive disciplinary climate is directly linked to high achievement in both the sophomore and senior years, controlling for influential student background characteristics such as race/ethnicity and socioeconomic status.[31] Chubb and Moe (1990) identified an indirect relationship that operates on student achievement through student perception of disciplinary fairness or effectiveness. Students in schools with greater average gains in student achievement view the discipline in their schools as:

> Fairer and more effective than the students in the schools at the bottom of the distribution view the discipline in theirs.... The important observation is that teaching in high-performance schools apparently takes place in a more disciplined atmosphere requiring less effort by teachers to maintain order. How this discipline is maintained—whether it benefits from the team-like organization of successful schools or relies on tough sanctions—is not clear. But what is clear is that the classes of high performance schools are subject to fewer intrusions and disruptions of all kinds ... the product is more class time for the educational activities deemed by teachers to be most appropriate (p. 98–99).

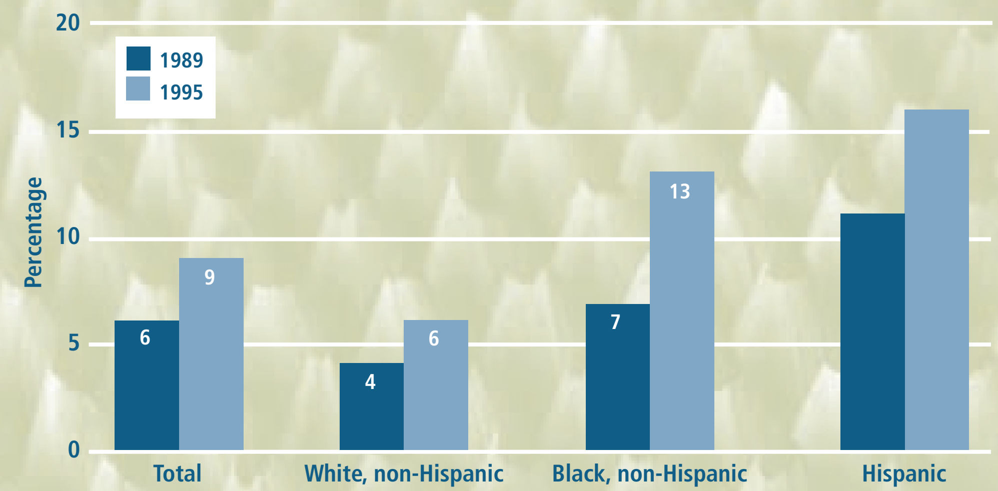
Violence in the school and immediate neighborhood may also reduce the likelihood of graduating from high school (Grogger 1997). Controlling for several background characteristics such as race/ethnicity and socioeconomic status, students in schools with substantial violence had lower math scores by 0.20 of a standard deviation and are 5.7 percentage points less likely to graduate than students in schools with minimal violence. Indeed, some student offenses are likely to affect student achievement negatively in every subject area. Barton, Coley, and Wenglinsky (1998) found that both serious and nonserious offenses (student tardiness, fighting, suspensions, and arrests) are negatively related to achievement gains between 8th and 12th grades in mathematics, reading, social studies, and science.

Serious violent crime incidents were reported in 10 percent of all public schools in 1996–97 (Kaufman et al. 1999). The level of school-related criminal behavior has changed little between 1976 and 1997, and no differences in victimization rates were found between white and black high school seniors in 1997 (Wirt et al. 1999). Figure 4.1 shows, however, that the percentage of middle and high school students who fear attack or other bodily harm while at school has been on the rise. In each year, a larger proportion of black and Hispanic students than white students feared attacks at school, and the percentage of black students who feared for their safety nearly doubled from 1989 through 1995 (Kaufman et al. 1999).

Current national data present a comprehensive picture of discipline incidents and identify the prevalence of various school policies related to safety and discipline, but they are

[31] Bryk, Lee, and Holland (1993) use the High School and Beyond data set to identify an approach that seems to pay off in a measurably low incidence of disruptive behavior and low rates of student disengagement, at least in the Catholic schools studied. The approach common to those Catholic schools combines a formal philosophy statement with unplanned daily encounters between faculty and students. It is during these otherwise routine encounters, they say, that Catholic school faculty take the opportunity to instill in students a sense of personal responsibility and commitment to appropriate behavior. Learning is facilitated, they say, when Catholic school students honor and respond to the commitments that parents, teachers, and others have made to giving them an opportunity to learn.

Figure 4.1 Percentage of students ages 12 through 19 who reported fearing being attacked or harmed at school, by race-ethnicity: 1989 and 1995



NOTE: Includes students who reported that they sometimes or most of the time feared being victimized in this way. "At school" means in the school building, on the school grounds, or on a school bus.

SOURCE: U.S. Department of Justice, Bureau of Justice Statistics, School Crime Supplement to the National Crime Victimization Survey, 1989 and 1995.

limited in their usefulness because the data do not include information that research suggests is relevant in determining the success of discipline policies, such as how the policy is implemented and its perceived fairness among the school community (Bryk, Lee, and Holland 1993; Chubb and Moe 1990).

E. INDICATOR 13: ACADEMIC ENVIRONMENT

One of the most frequently cited attributes of schools with high levels of student learning and consistent positive student growth is that schools emphasize (and students take advantage of) a challenging and appropriate curriculum (Bryk, Lee, and Holland 1993; Chubb and Moe 1990). Unless students are productively engaged in appropriately challenging academic study, a school's orderly atmosphere will have little effect on student learning. Knowing what school-level policies and schoolwide actions encourage students into recommended course-taking patterns and toward academic success is useful information for schools searching for greater success. To understand these issues, research suggests that we must know whether and how student performance is related to the school's course offerings (academic versus vocational) and the students' responses to those course offerings and the extent to which the school encourages students into particular enrollment patterns and holds high expectations for their success (Bryk, Lee, and Holland 1993; Chubb and Moe 1990).

National information on the influence of coursetaking and the academic environment on achievement primarily derives from the High School and Beyond and the National Education Longitudinal Study of 1988. Results have been widely reported in several analyses. Chubb and Moe (1990) use the High School and Beyond longitudinal data to suggest that "academic program participation has a strong, independent effect on achievement gains.... All things being equal, academic programs promote academic achievement" (p. 131). Bryk, Lee, and Holland (1993), using the same data, concur. Controlling for student characteristics, they find that student achievement is greater in schools where, among other factors, students take a higher number of math courses.

Does this suggest that a school can raise its student achievement levels merely by offering students a narrow, highly academic program with limited course selection? Not entirely. Bryk, Lee, and Holland (1993) suggest that taking more math and science courses and enrolling in honors programs are strongly related to achievement in all content areas in part because "more able students generally take advanced courses in all academic areas" (pp. 205–207).

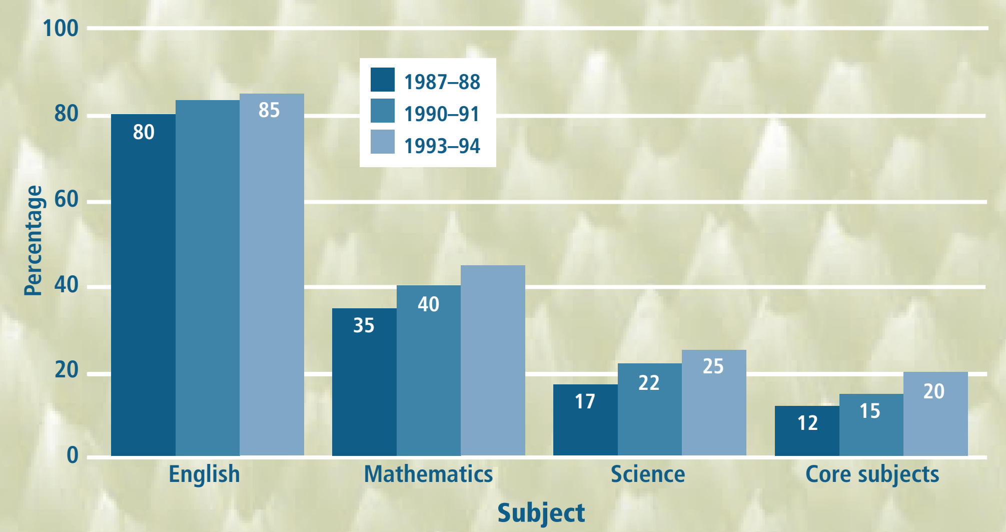
If course offerings contribute to achievement gains, it will be useful to understand the mechanism through which that occurs and the expected magnitude. Chubb and Moe (1990) suggest gains are related to the academic environment in the following way. The percentage of students enrolled in the academic track is an "indicator of the programmatic orientation of the school [that] gauges the emphasis that the school places on academic work [and] captures how aggressively schools track students into academic programs and how extensively academic work contributes to the school's general climate" (p. 131). The effect on student achievement is not insignificant. Controlling for student characteristics such as socioeconomic status, they estimate that differences in tracking policies across effective and ineffective schools amount to 0.13 of a standard deviation difference in achievement each year over the final two years of high school, equal to approximately 30 percent of the total influence of school organization on achievement.[32]

More recent research into the academic environment and its effects on student achievement has tried to distinguish between differences in learning resulting from two currently prevailing characterizations of school organization and climate: academic press and communitarian organization. Schools emphasizing academic press typically implement strong achievement-related school goals that are manifested in shared high expectations for student achievement. Communitarian schools, on the other hand, have communal or shared forms of organization and are usually engaged in reorganizing instruction, altering authority, personalizing relationships within the school, and developing a sense of community. Research on the relative effects of each has been inconclusive so far. Some researchers found that academic press was linked to achievement (Phillips 1997; Shouse 1996), while others found no difference (Lee and Smith 1995).

It is perhaps easier to measure academic press than the more multidimensional concept of communitarianism. Communitarianism is especially prone to the data concerns identified at the beginning of this chapter because it represents an organizational ethos, not a set of policies. Thus, measuring it requires overcoming the difficulties associated

[32]Chubb and Moe suggest that the effect on student achievement that may result from academic tracking could depend on how researchers specify their statistical models. For two reasons, they say, it is inappropriate to use academic program participation as a factor to assess the effects of school organization. First, since enrolling in an academic program is partly a function of the ability and background of the student, it is not appropriate to assign those effects only to the school. Second, as an estimate of the probability that any given student is academically tracked, the measure also operates as a student-level proxy that gauges the influence of course work on individual gains in achievement. Instead, Chubb and Moe suggest the appropriate model should include the individual effect of student academic coursetaking (by creating a dichotomous variable measured at the student level) and the school-level effect on student learning (derived from the school's influence on the students' choice of study programs).

Figure 4.2 Percentage of public school districts with graduation requirements at or above NCEE recommendations, by subject: School years 1987–88, 1990–91, 1993–94



NOTE: The National Commission on Excellence in Education (NCEE) recommends that students complete 4 years of English, 3 years of mathematics, 3 years of science, and 3 years of social studies. Core subjects include English, mathematics, science, and social studies. Data for social studies are not shown separately.

SOURCE: U.S. Department of Education, National Center for Education Statistics, *The Condition of Education 1998,* Schools and Staffing Survey.

with creating numerical representations of human incentives, motivations, interactions, and opinions. In addition, currently available data of this sort come from only a small and perhaps unrepresentative sample of teachers within schools.

Academic press is more tangible and can be represented by student course offerings and requirements. For example, Figure 4.2 shows how academic expectations have changed over time. The figure shows that the percentage of public school districts with graduation requirements that meet or exceed the National Commission on Excellence in Education recommendations (four years of English, three years of mathematics, three years of science, three years of social studies, and a half a year of computer science) increased from 12 to 20 percent between 1987–88 and 1993–94, and that districts were more likely to meet or exceed those recommendations for English than for mathematics or science (Wirt et al. 1998). In addition, students appear to have responded to these policy changes. The percentage of graduates completing the NCEE requirements in English, science, mathematics, and social studies tripled between 1982 and 1994 (Center on Education Policy and American Youth Policy Forum 2000). A common criticism of the NCEE recommendations is that they specify only the number of courses to be taken, not their rigor. But, as noted in the previous chapter, there is evidence that students have been enrolling in larger numbers in more difficult courses. From 1982 to 1998, there was an increase in the percentage of students enrolling in higher-level mathematics and science courses such as algebra II, trigonometry, precalculus, calculus, physics I, chemistry I, chemistry II, or physics II (Wirt et al. 2000).

For the reasons detailed in Chapter III, course titles provide only a rough approximation of content (Porter et al. 1993; Shouse 1996). But given the estimated effect size on student achievement resulting from academic organization, we need to continue the push for better identifying organizational nuances that might unlock our understanding of how organizational philosophies encourage students to excel and promote student learning.

F. SUMMARY

How schools approach educational leadership and school goals, develop a professional community, and establish a climate that minimizes discipline problems and encourages academic excellence clearly affect school quality and student learning. For three reasons, however, the effect of school-level characteristics are more difficult to ascertain than those of teachers and classrooms described in the previous two chapters. First, even though they are integral to a school, these characteristics are difficult to define and measure. Second, their effect on student learning is likely to be exerted indirectly through teachers and classrooms, compounding the measurement problem. And last, with some exceptions, reliable school-representative information about these indicators of quality is minimal. These difficulties should not overshadow the importance of collecting such data to learn more about how these characteristics operate and affect student learning through teachers and classrooms. The preponderance of national, regional, and local efforts to develop quality schools heightens the benefits that would be derived from additional and more refined and reliable school-representative measures of school characteristics.

V. CONCLUSION

Thirteen indicators of school quality that recent research suggests are related to improving student learning are the focus of this report. These indicators of quality cover three areas of schools: the characteristics of teachers, classrooms, and schools as organizations. Many of the indicators in this report are supported by evidence that a statistical association exists between the indicator and student learning. But in several instances, even though experts would agree that certain indicators should enhance student learning, there is not always concrete statistical evidence of an association (examples include professional development, school leadership, school goals, and a school's professional community). Statistical evidence may be lacking because the indicator data need to be improved.

This chapter has four purposes: (1) to synthesize the report's findings regarding the caliber of the data on the indicators related to teachers, classrooms, and schools; (2) to explain why some indicators are higher caliber than others; (3) to highlight which indicators need improving; and (4) to summarize the major findings regarding the current status of school quality on each dimension of schooling for which data are adequate.

A. QUALITY OF THE DATA

To assess the status of the data on each indicator of quality, we searched the national data to determine which data best represented each indicator and then interpreted the status of the 13 dimensions of schools using these data. Interpreting the data on measures of some indicators (such as teacher assignment) was relatively straightforward; interpreting data for other indicators (such as leadership) was less so. The quality of the data depends on one of at least two factors: (1) the complexity of the characteristics being measured and (2) whether the dimension of schooling has been studied using national data for an extended period of time.

The quality of data on these indicators varies in predictable ways. Where the dimension being measured is straightforward (teaching assignment, teacher experience, and class size) or has been measured for an extended period of time (the standardized tests used to gauge teacher academic skills), the data are high quality. Where little information exists about a particularly important facet of an indicator (professional development, course content, technology, discipline, and academic environment), the quality of the data is somewhat compromised. And where an indicator is more complex than are measures of a phenomenon (pedagogy, leadership, goals, and professional community), data quality is poor. Data quality designations for each of the 13 indicators are listed in Table 5.1.

The indicators of teacher assignment, teacher experience, and class size each represent straightforward concepts and are easy to measure, and the data are high quality. In addition, data on teacher experience and class size have been collected for several decades, further ensuring the quality of the data. Data on teacher academic skills are also high quality but less straightforward. While the academic skills of teachers constitute only one aspect of teaching ability, standardized tests that measure the academic skills of future teachers are quite advanced and have consistently shown links to student learning.

Data on indicators of professional development, course content, technology, discipline, and academic environment are of moderate quality. The national data-collection efforts of these specific measures are relatively new compared with the efforts to collect data on high-quality indicators, and the indicators are more complex than the data currently collected. Consequently, data on professional development are limited in key respects and provide little insight into important principles of successful programs (for example, current data tell us mostly about the prevalence of professional development, but not much about its structure and quality). National data on indicators of course content and academic environment are based primarily on course titles, and consequently their usefulness is hampered by their being too vague. Similarly, current data on technology primarily measure the availability of hardware and access to the Internet and provide too little information on the instructional role of technology in the classroom. Nationally representative data on school discipline incidents and on school discipline policies are well defined and of good quality, but there are limited data documenting this relationship to student learning, the implementation of discipline policies, and their perceived fairness.

Only poor-quality data are available on pedagogy, leadership, school goals, and professional community. These indicators are complex to measure and historically have not been prominent parts of national data-collection efforts. The teaching process consists of a complex set of interactions between students, the teacher, and the curriculum, making it hard to isolate and measure critical elements of pedagogy. Measuring human actions, incentives, and opinions to estimate the effects of school-level attributes such as leadership, goals, and professional community is an equally complex task.

As a group, the teacher-focused measures of school quality are less complex and have been collected for some time. School-level attributes of quality are nearly the opposite. Indicators with lower-quality data provide an incentive and direction for improved national data collection.

But even the high-quality measures are sometimes compromised by not being used or defined properly and by not being kept up to date. How a measure is used or defined can alter our perception of school quality. As discussed in the teacher and classrooms chapters, teacher assignment and class size can each be defined in substantively different ways, and the different definitions provide different pictures of the magnitude of the out-of-field teaching problem and the average size of classrooms.

Furthermore, the value of measures depreciates if the data are old. Some of the high-quality measures are not collected frequently. A national measure of out-of-field teaching using the "any-mismatch" approach is obtained every four years. Other high-

Table 5.1 Quality of national school quality indicator data

High Quality	Moderate Quality	Poor Quality
Teacher Assignment	Professional Development	Pedagogy
Teacher Experience	Technology	Goals
Teacher Academic Skills	Course Content	School Leadership
Class Size	Discipline	Professional Community
	Academic Environment	

quality measures are not collected routinely. There are no regularly scheduled efforts to use standardized tests to measure the academic skills of prospective teachers or to monitor the course-taking patterns of students.

To monitor the status of school quality accurately, better measures are needed, especially for indicators of pedagogy, leadership, goals, and professional community. Furthermore, certain important facets of professional development, course content, technology, academic environment, and discipline are missing. Finally, even when quality data are available, they lose their value if they are not appropriately defined and kept up to date. Moreover, even though experts would agree that certain indicators should enhance student learning, there is not always concrete statistical evidence to support their supposition; improving the data collected on the dimensions of schools thought to be associated with school quality should help us better understand the relationship of these indicators to student learning.

B. THE STATUS OF SCHOOL QUALITY

Teacher indicators show that the teaching profession attracts those with low academic skills who have been entering the profession in much higher numbers than teachers with high academic skills (Henke, Chen, and Geis 2000). Inexperienced teachers and those with low academic skills are more likely to teach in high-poverty schools and schools with high concentrations of minority students (Henke, Chen, and Geis 2000). Although new-teacher induction programs are on the rise, continuing teachers rarely participate in staff renewal and enhancement activities for more than one day each year (Choy and Ross 1998). The classroom indicators show more positive trends. Students are taking more rigorous academic courses in high school, and more computers are available and used more frequently (Smith et al. 1997; Wirt et al. 2000). However, minority high school graduates were usually less likely than Asian/Pacific Islander or non-Hispanic white students to enroll in these classes (Wirt et al. 2000). In addition, students in high-poverty schools have less classroom access to the Internet (U.S. Department of Education 2000b). Although the average public elementary school class in 1998 had 23 students, the strongest gains in student achievement seem to occur in classes with 13 to 20 students, especially for disadvantaged and minority students (Krueger 1998; Mosteller, Light, and Sachs 1996; Robinson and Wittebols 1986). School indicators show that schools have been raising their academic expectations and that an increased number of students are enrolling in algebra I, geometry, algebra II, biology, chemistry, and physics (Smith et al. 1997). However, serious violent crime incidents occurred in 10 percent of all public schools in 1996–97, and the percentage of middle and high school students who fear attack or other bodily harm while at school has been on the rise (Kaufman et al. 1999).

The findings documented in this report, like all research, are time sensitive and part of an iterative process. The status of schools as identified by indicators with quality data is changing rapidly and will need to be continually updated. As research on school effectiveness proceeds, indicators with only poor-quality data will need to be improved to understand the complete picture of school quality as called for by the Special Study Panel on Education Indicators for the National Center for Education Statistics.

REFERENCES

1999 National Education Summit. (1999). *1999 Action Statement.* 1999 National Education Summit. Available: *http://www.summit99.org/press/actionstatement.html.*

Achilles, C. M. (1996). Response to Eric Hanushek: Students Achieve More in Smaller Classes. *Educational Leadership,* 76–77.

Anderson, R. E., and A. Ronnkvist. (1999). *The Presence of Computers in American Schools.* Irvine, CA: Center for Research on Information Technology and Organizations.

Ballou, D. (1996). Do Public Schools Hire the Best Applicants? *The Quarterly Journal of Economics,* 97–133.

Barron's. (1984). *Profiles of American Colleges.* Woodbury, NY: Barron's Educational Series.

Barton, P. E., R. J. Coley, and H. Wenglinsky. (1998). *Order in the Classroom: Violence, Discipline, and Student Achievement.* Princeton, NJ: Policy Information Center, Educational Testing Service.

Becker, H. J., J. L. Ravitz, and Y. T. Wong. (1999). *Teacher and Teacher-Directed Student Use of Computers and Software.* Irvine, CA: Center for Research on Information Technology and Organizations.

Bobbitt, S. A., and M. M. McMillen. (1994). *Qualifications of the Public School Teacher Workforce: 1988 and 1991* (NCES 95–665). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Bohrnstedt, G. W., and B. M. Stecher. (1999). *Class Size Reduction in California: Early Evaluation Findings, 1996–1998 (CSR Research Consortium, Year 1 Evaluation Report).* Palo Alto, CA: American Institutes for Research.

Brophy, J., and T. L. Good. (1986). *Teacher Behavior and Student Achievement.* In *Handbook of Research on Teaching,* edited by M. Wittrock. New York: MacMillan.

Bruschi, B. A., and R. J. Coley. (1999). *How Teachers Compare: The Prose, Document, and Quantitative Skills of America's Teachers.* Princeton, NJ: Policy Information Center, Educational Testing Service.

Bryk, A. S., V. E. Lee, and P. B. Holland. (1993). *Catholic Schools and the Common Good.* Cambridge, MA and London: Harvard University Press.

Bryk, A. S., and Y. M. Thum. (1989). The Effects of High School Organization on Dropping Out: An Exploratory Investigation. *American Educational Research Journal, 26*(3): 353–383.

Burkam, D. T., V. E. Lee, and B. A. Smerdon. (1997). *Mathematics, Foreign Language, and Science Coursetaking and the NELS: 88 Transcript Data.* Ann Arbor: University of Michigan.

Burstein, L., L. M. McDonnell, J. Van Winkle, T. Ormseth, J. Mirocha, and G. Guitton. (1995). *Validating National Curriculum Indicators.* Santa Monica, CA: RAND Corporation.

Case, R., and C. Bereiter. (1984). From Behaviorism to Cognitive Development. *Instructional Science, 13:* 141–158.

Center on Education Policy and American Youth Policy Forum. (2000). *Do You Know the Good News about American Education?* Washington, DC: Center on Education Policy and American Youth Policy Forum.

Choy, S. P., and M. Ross. (1998). *Toward Better Teaching: Professional Development in 1993–94* (NCES 98–230). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Chubb, J. E., and T. M. Moe. (1990). *Politics, Markets, and America's Schools.* Washington, DC: The Brookings Institution.

Cobb, R., and L. P. Steffe. (1983). The Constructivist Researcher as Teacher and Model Builder. *Journal for Research in Mathematics Education, 14:* 83–94.

Cohen, D. K., and H. C. Hill. (2000). Instructional Policy and Classroom Performance: The Mathematics Reform in California. *Teachers College Record, 102*(2): 294–343.

Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York. (1966). *Equality of Educational Opportunity.* Washington, DC: U.S. Government Printing Office.

Coley, R. J., J. Cradler, and P. K. Engel. (1997). *Computers and Classrooms: The Status of Technology in U.S. Schools.* Princeton, NJ: Policy Information Center, Educational Testing Service.

CPRE Policy Brief. (1995). *Helping Teachers Teach Well: Transforming Professional Development.* New Brunswick, NJ: CPRE.

Darling-Hammond, L. (1998). Teachers and Teaching: Testing Policy Hypotheses from a National Commission Report. *Educational Researcher, 27*(1): 5–15.

Darling-Hammond, L. (2000). Teacher Quality and Student Achievement: A Review of State Policy Evidence. *Education Policy Analysis Archives, 8*(1). Available: *http://olam.ed.asu.edu/epaa/v8n1/.*

Davis, R. B. (1984). *Learning Mathematics: The Cognitive Science Approach to Mathematics Education.* Norwood, MA: Ablex.

Deal, T. E., and K. D. Peterson. (1998). *Shaping School Culture: The Heart of Leadership.* San Francisco: CA: Jossey-Bass.

Decker, P. T., J. K. Rice, and M. T. Moore. (1997). *Education and the Economy: An Indicators Report* (NCES 97–269). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Education Week. (1998). Quality Counts. *Supplement to Education Week, 17.*

Ehrenberg, R. G., and D. Brewer. (1994). Do School and Teacher Characteristics Matter? Evidence from High School and Beyond. *Economics of Education Review, 13*(1): 1–17.

Ehrenberg, R. G., and D. J. Brewer. (1995). Did Teachers' Verbal Ability and Race Matter in the 1960s? Coleman Revisited. *Economics of Education Review, 14*(1): 1–21.

Ferguson, R. F. (1991). Paying for Public Education: New Evidence on How and Why Money Matters. *Harvard Journal on Legislation, 28*(2): 465–499.

Ferguson, R. F. (1998). *Can Schools Narrow the Black-White Test Score Gap?* In *The Black-White Test Score Gap,* edited by C. Jencks , and M. Phillips. Washington, DC: Brookings Institution Press.

Ferguson, R. F., and H. Ladd. (1996). *How and Why Money Matters: An Analysis of Alabama Schools.* In *Holding Schools Accountable: Performance-Based Reform in Education,* edited by H. F. Ladd. Washington, DC: Brookings Institution.

Finn, J. D. (1998). *Class Size and Students at Risk: What is Known? What is Next?* U.S. Department of Education. Washington, DC: Office of Educational Research and Improvement.

Fiore, T. A., T. R. Curtin, and C. H. Hammer. (1997). *Public and Private School Principals in the United States: A Statistical Profile, 1987–88 and 1993–94* (NCES 97–455). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Fullan, M. (1998). Leadership for the 21st Century: Breaking the Bonds of Dependency. *Educational Leadership, 55*(7): 6–10.

Gitomer, D. H., A. S. Latham, and R. Ziomek. (1999). *The Academic Quality of Prospective Teachers: The Impact of Admissions and Licensure Testing.* Princeton, NJ: Teaching and Learning Division, Educational Testing Service.

Glass, G. V., and M. L. Smith. (1979). Meta-Analysis of Research on Class Size and Achievement. *Education Evaluation and Policy Analysis, 1*(1): 2–16.

Glennan, T. K., and A. Melmed. (1996). *Fostering the Use of Educational Technology: Elements of a National Strategy.* Santa Monica, CA: RAND.

Goals 2000. (1999). *Building Bridges: The Mission and Principles of Professional Development.* Washington, DC: U.S. Department of Education. Available: *http://www.ed.gov/G2K/bridge.html.*

Goldhaber, D. D., and D. J. Brewer. (1997). Evaluating the Effect of Teacher Degree Level on Educational Performance. In W. Fowler (Ed.), *Developments in School Finance, 1996* (NCES 97– 535) (pp. 197–210). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Greenwald, R., L. V. Hedges, and R. D. Laine. (1996). The Effect of School Resources on Student Achievement. *Review of Educational Research, 66*(3): 361–396.

Grissmer, D., A. Flanagan, and S. Williamson. (1998). *Why Did the Black-White Score Gap Narrow in the 1970s and 1980s?* In *The Black-White Test Score Gap,* edited by C. Jencks , and M. Phillips. Washington, DC: Brookings Institution Press.

Grogger, J. (1997). Local Violence and Educational Attainment. *The Journal of Human Resources, 32*(4): 659–682.

Haney, W., G. Madaus, and A. Kreitzer. (1987). *Charms Talismanic: Testing Teachers for the Improvement of American Education.* In *Review of Research on Education,* edited by E. Rothkopf. Washington, DC: American Educational Research Association.

Hanushek, E. A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature, 24*(3): 1141–1177.

Hanushek, E. A. (1989). The Impact of Differential Expenditures on School Performance. *Educational Researcher, 18*(4): 45–51.

Hanushek, E. A. (1992). The Trade-Off between Child Quantity and Quality. *Journal of Political Economy, 100*(1): 84–117.

Hanushek, E. A. (1996). A More Complete Picture of School Resource Policies. *Review of Educational Research, 66*(3): 397–409.

Hanushek, E. A. (1998). *The Evidence on Class Size.* Rochester, NY: University of Rochester, W. Allen Wallis Institute of Political Economy.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin. (1998). *Teachers, Schools and Academic Achievement.* National Bureau of Economic Research Working Papers. Available: *http://papers.nber.org/papers/W6691.*

Hanushek, E. A., J. F. Kain, and S. G. Rivkin. (1999). *Do Higher Salaries Buy Better Teachers?* National Bureau of Economic Research Working Papers. Available: *http://www.nber.org/papers/w7082.*

Hawkins, E. F., F. B. Stancavage, and J. A. Dossey. (1998). *School Policies and Practices Affecting Instruction in Mathematics: Findings from the National Assessment of Educational Progress* (NCES 98–495). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Hedges, L. V., R. D. Laine, and R. Greenwald. (1994). Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes. *Educational Researcher, 23*(3): 5–14.

Henke, R. R., X. Chen, and S. Geis. (2000). *Progress through the Teacher Pipeline: 1992–93 College Graduates and Elementary/Secondary School Teaching as of 1997* (NCES 2000–152). U.S. Department of Education. Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement.

Henke, R. R., X. Chen, and G. Goldman. (1999). *What Happens in Classrooms? Instructional Practices in Elementary and Secondary Schools, 1994–95* (NCES 1999–348). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Henke, R. R., S. P. Choy, X. Chen, S. Geis, and M. N. Alt. (1997). *America's Teachers: Profile of a Profession, 1993–94* (NCES 97–460). U.S. Department of Education, National Center for Education Statistics, Office of Educational Research and Improvement. Washington, DC: U.S. Government Printing Office.

Henke, R. R., S. P. Choy, S. Geis, and S. P. Broughmann. (1996). S*chools and Staffing in the United States: A Statistical Profile, 1993–94* (NCES 96–124). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Henke, R. R., S. Geis, and J. Giambattista. (1996). *Out of the Lecture Hall and into the Classroom: 1992–93 College Graduates and Elementary/Secondary School Teaching* (NCES 96–899). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Heyns, B. (1988). Educational Defectors: A First Look at Teachers Attrition in the NLS–72. *Educational Researcher, 17*(3): 24–32.

Hiebert, J., and D. Wearne. (1993). Instructional Tasks, Classroom Discourse, and Students' Learning in Second-Grade Arithmetic. *American Educational Research Journal, 30:* 393–425.

Ingersoll, R. M. (1999). The Problem of Underqualified Teachers in American Secondary Schools. *Educational Researcher, 28*(2): 26–37.

Kaufman, P., X. Chen, S. P. Choy, S. A. Ruddy, A. Miller, K. A. Chandler, C. D. Chapman, M. R. Rand, and P. Klaus. (1999). *Indicators of School Crime and Safety, 1999* (NCES 1999–057). U.S. Departments of Education and Justice. Washington, DC: National Center for Education Statistics.

King, M. B., K. S. Louis, H. M. Marks, and K. D. Peterson. (1996). *Participatory Decision Making.* In *Authentic Achievement: Restructuring Schools for Intellectual Quality,* edited by Fred M. Newmann and Associates. San Francisco: Jossey-Bass.

Krueger, A. B. (1998). *Experimental Estimates of Education Production Functions.* Princeton, NJ: Princeton University Industrial Relations Section.

Lampert, M. (1986). Knowing, Doing and Teaching Mathematics. *Cognition and Instruction, 3:* 305–342.

Lee, V. E., and J. B. Smith. (1995). Effects of High School Restructuring and Size on Early Gains in Achievement and Engagement. *Sociology of Education, 68:* 241–270.

Lee, V. E., and J. B. Smith. (1996). Collective Responsibility for Learning and Its Effects on Gains in Achievement for Early Secondary School Students. *American Journal of Education, 104:* 103–146.

Leithwood, K., D. Jantzi, and A. Fernandez. (1995). *Transformational Leadership and Teachers' Commitment to Change.* In *Reshaping the Principalship: Insights from Transformational Reform Efforts,* edited by J. Murphy , and K. Louis. Thousand Oaks, CA: Corwin Press.

Lesh, R., and M. Landau (Eds.). (1983) *Acquisition of Mathematics Concepts and Processes.* New York: Academic Press.

Levine, D. U., and L. W. Lezotte. (1990). *Unusually Effective Schools: A Review and Analysis of Research and Practice.* Madison, WI: National Center for Effective Schools Research and Development.

Lewis, L., B. Parsad, N. Carey, N. Bartfai, E. Farris, and B. Smerdon. (1999). *Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers* (NCES 1999–080). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Lipsey, M. W., and D. B. Wilson. (1993). The Efficacy of Psychological, Educational, and Behavioral Treatment. *American Psychologist, 48*(12): 1181–1209.

Louis, K. S., S. D. Kruse, and Associates. (1995). *Professionalism and Community: Perspectives on Reforming Urban Schools.* Thousand Oaks, CA: Corwin Press.

Louis, K. S., S. D. Kruse, and H. M. Marks. (1996). *Schoolwide Professional Community.* In *Authentic Achievement: Restructuring Schools for Intellectual Quality,* edited by Fred M. Newmann and Associates. San Francisco: Jossey-Bass.

Marks, H. M., K. B. Doane, and W. B. Secada. (1996). *Support for Student Achievement.* In *Authentic Achievement: Restructuring Schools for Intellectual Quality,* edited by Fred M. Newmann and Associates. San Francisco: Jossey-Bass.

Mayer, D. P. (1998). Do New Teaching Standards Undermine Performance on Old Tests? *Educational Evaluation and Policy Analysis, 20*(2): 53–73.

Mayer, D. P. (1999a). Invited Commentary: Moving toward Better Instructional Practice Data. *Education Statistics Quarterly, 1*(2): 17–20.

Mayer, D. P. (1999b). Measuring Instructional Practice: Can Policy Makers Trust Survey Data? *Educational Evaluation and Policy Analysis, 21*(1): 29–45.

McKnight, C. C., F. J. Crosswhite, J. A. Dossey, E. Kifer, J. O. Swafford, K. J. Travers, and T. J. Cooney. (1987). *The Underachieving Curriculum: Assessing U.S. School Mathematics From an International Perspective.* Champaign, IL: Stipes Publishing Company.

Monk, D. H., and J. King. (1994). *Multi-level Teacher Resource Effects on Pupil Performance in Secondary Mathematics and Science: The Role of Teacher Subject Matter Preparation.* In *Contemporary Policy Issues: Choices and Consequences in Education,* edited by R. Ehrenberg. Ithaca, NY: ILR Press.

Mosteller, F. (1995). The Tennessee Study of Class Size in the Early Grades. *The Future of Children, 5*(2): 113–127.

Mosteller, F., R. J. Light, and J. A. Sachs. (1996). Sustained Inquiry in Education: Lessons from Skill Grouping and Class Size. *Harvard Education Review, 66*(4): 797–842.

Mosteller, F., and D. P. Moynihan (Eds.). (1972) *On Equality of Educational Opportunity.* New York: Random House.

Mullens, J., and K. Gayler. (1999). *Measuring Classroom Instructional Processes: Using Survey and Case Study Field Test Results to Improve Item Construction* (NCES 1999–08). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Mullens, J. E., M. S. Leighton, K. G. Laguarda, and E. O'Brien. (1996). *Student Learning, Teacher Quality, and Professional Development: Theoretical Linkages, Current Measurement, and Recommendations for Future Data Collection* (NCES 96–28). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Murnane, R. J., and R. Levy. (1996a). *Evidence from 15 Schools in Austin, Texas.* In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success,* edited by G. Burtless. Washington, DC: Brookings Institution.

Murnane, R. J., and R. Levy. (1996b). *Teaching the New Basic Skills: Principles for Educating Children to Thrive in a Changing Economy.* New York: The Free Press.

Murnane, R. J., and R. Olsen. (1990). The Effects of Salaries and Opportunity Costs on Length of Stay in Teaching. *The Journal of Human Resources, 21*(1): 106–124.

Murnane, R. J., and B. R. Phillips. (1981). Learning by Doing, Vintage, and Selection: Three Pieces of the Puzzle Relating Teaching Experience and Teaching Performance. *Economics of Education Review, 1*(4): 453–465.

Murnane, R. J., and S. A. Raizen (Eds.). (1988) *Improving Indicators of the Quality of Science and Mathematics Education in Grades K-12.* Washington, DC: National Academy Press.

Murnane, R. J., J. D. Singer, J. B. Willett, J. J. Kemple, and R. J. Olson. (1991). *Who Will Teach? Policies That Matter.* Cambridge, MA: Harvard University Press.

Murphy, J. (1994). *Transformational Change and the Evolving Role of the Principal: Early Empirical Evidence.* In *Reshaping the Principalship: Insights from Transformational Reform Efforts,* edited by J. Murphy , and K. Louis. Thousand Oaks, CA: Corwin Press.

National Commission on Teaching and America's Future. (1996). *What Matters Most: Teaching for America's Future.* New York: The National Commission on Teaching and America's Future.

National Council of Teachers of Mathematics. (1991). *Professional Standards for Teaching Mathematics.* Reston, VA: National Council of Teachers of Mathematics.

National Education Goals Panel. (1995). *Teacher Education and Professional Development: Report of the Goal 4 Resource Group.* Washington, DC: National Education Goals Panel.

National Foundation for the Improvement of Education. (1996). *Teachers Take Charge of Their Learning: Transforming Professional Development for Student Success.* Washington, DC: National Foundation for the Improvement of Education.

Newmann, F. M., and G. G. Wehlage. (1995). *Successful School Restructuring: A Report to the Public and Educators by the Center on Organization and Restructuring of Schools.* Madison, WI: Center on Organization and Restructuring of Schools, University Wisconsin-Madison School of Education, Wisconsin Center for Education Research.

Odden, A. (1990). Class Size and Student Achievement: Research-Based Policy Alternatives. *Educational Evaluation and Policy Analysis, 12*(2): 213–227.

Phillips, M. (1997). What Makes Schools Effective? A Comparison of the Relationships of Communitarian Climate and Academic Climate to Mathematics Achievement and Attendance during Middle School. *American Educational Research Journal, 34*(4): 633–662.

Porter, A. C. (1991). Creating a System of School Process Indicators. *Educational Evaluation and Policy Analysis, 13*(1): 13–29.

Porter, A. C., M. W. Kirst, E. J. Osthoff, J. L. Smithson, and S. A. Schneider. (1993). *Reform Up Close: An Analysis of High School Mathematics and Science Classrooms.* Madison, WI: Center on Organization and Restructuring of Schools, University of Wisconsin-Madison, School of Education, Wisconsin Center for Education Research.

Powell, A. G., E. Farrar, and D. K. Cohen. (1985). *The Shopping Mall High School: Winners and Losers in the Educational Marketplace.* Boston, MA: Houghton-Mifflin.

President's Committee of Advisors on Science and Technology, Panel on Educational Technology. (1997). *Report to the President on the Use of Technology to Strengthen K-12 Education in the United States.* Washington, DC: The White House.

Purkey, S. C., and M. S. Smith. (1983). Effective Schools: A Review. *The Elementary School Journal, 83*(4): 427–452.

Raizen, S. A., and L. V. Jones. (1985). *Indicators of Precollege Education in Science and Mathematics: A Preliminary Review.* Washington, DC: National Academy Press.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain. *Teachers, Schools and Academic Achievement.* (1998) Paper presented at the Association for Public Policy Analysis and Management, New York City.

Robinson, G. E., and J. H. Wittebols. (1986). *Class Size Research: A Related Cluster Analysis for Decision Making.* Arlington, VA: Education Research Service.

Rollefson, M. R., and T. M. Smith. (1997). *Do Low Salaries Really Draw the Least Able into the Teaching Profession?* In *Research on the Education of Our Nation's Teachers: Teacher Education Yearbook V,* edited by D. M. Byrd , and D. J. McIntyre. Thousand Oaks, CA: Corwin Press.

Rotberg, I. C. (1998). Interpretation of International Test Score Comparisons. *Science, 280:* 1030–1031.

Sanders, W. L., and J. C. Rivers. (1996). *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement.* Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Sandham, J. L. (1998). Massachusetts Chief Resigns in Protest amid Test Flap. *Education Week, 17*(42): 20.

Schmidt, W. H., C. C. McKnight, L. S. Cogan, P. M. Jakwerth, and R. T. Houang. (1999). *Facing the Consequences: Using TIMSS for a Closer Look at U.S. Mathematics and Science Education.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt, W. H., C. C. McKnight, and S. A. Raizen. (1997). *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education.* Boston: Kluwer Academic Publishers.

Schoenfeld, A. H. (1987). *What's All the Fuss About Metacognition?* In *Cognitive Science and Mathematics Education,* edited by A. H. Schoenfeld. Hillside, NJ: Lawrence Erlbaum Associates.

Sebring, P. A. (1987). Consequences of Differential Amounts of High School Coursework: Will the New Graduation Requirements Help? E*ducational Evaluation and Policy Analysis, 9* (3): 257– 273.

Shapson, S. M., E. N. Wright, G. Eason, and J. Fitzgerald. (1980). An Experimental Study of Effects of Class Size. *American Educational Research Journal, 17:* 141–152.

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

Shields, P. M., C. E. Esch, D. C. Humphrey, V. M. Young, M. Gaston, and H. Hunt. (1999). *The Status of the Teaching Profession: Research Findings and Policy Recommendations.* Santa Cruz, CA: The Center for the Future of Teaching and Learning.

Shouse, R. C. (1996). Academic Press and Sense of Community: Conflict, Congruence, and Implications for Student Achievement. *Social Psychology of Education, 1:* 47–68.

Slavin, R. E. (Ed.). (1989) *School and Classroom Organization.* Hillside, NJ: Lawrence Erlbaum Associates.

Smerdon, B., and S. Cronen. (2000). *Teachers' Tools for the 21st Century: A Report on Teachers' Use of Technology* (NCES 2000–102). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Smith, T. M., B. A. Young, Y. Bae, S. P. Choy, and N. Alsalam. (1997). *The Condition of Education 1997* (NCES 97–388). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Smylie, M. A. (1996). From Bureaucratic Control to Building Human Capital: The Importance of Teacher Learning in Education Reform. *Educational Researcher, 25*(9): 9–11.

Sterling, W. (1998). What Is the Massachusetts Teacher Exam Really Testing? *Education Week, 18*(15): 37.

Stigler, J. W., P. Gonzales, T. Kawanaka, S. Knoll, and A. Serrano. (1999). *The TIMSS Videotape Classroom Study: Methods and Findings from an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan, and the United States* (NCES 1999–074). U.S. Department of Education. Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement.

U.S. Department of Education. (1996). *Getting America's Students Ready for the 21st Century: Meeting the Technology Literacy Challenge.* Washington, DC.

U.S. Department of Education. (1998). *Reducing Class Size: What Do We Know?* Washington, DC.

U.S. Department of Education. (1999a). *Designing Effective Professional Development: Lessons from the Eisenhower Program.* Washington, DC.

U.S. Department of Education. (1999b). *The Initial Report of the Secretary on the Quality of Teacher Preparation.* Washington, DC.

U.S. Department of Education. (2000a). *Computer and Internet Access in Private Schools and Classrooms: 1995 and 1998* (NCES 2000–044). Washington, DC: National Center for Education Statistics.

U.S. Department of Education. (2000b). *Internet Access in Public Schools and Classrooms: 1994–99* (NCES 2000–086). Washington, DC: National Center for Education Statistics.

U.S. Department of Education, National Center for Education Statistics, Special Study Panel on Education Indicators for the National Center for Education Statistics. (1991). *Education Counts* (NCES 91–634) Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics, Office of Educational Research and Improvement. (1995). *A Profile of the American High School Senior in 1992* (NCES 95–384) Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics. (1997). *Pursuing Excellence: A Study of U.S. Fourth-Grade Mathematics and Science Achievement in International Context.* (NCES 97–255) Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, Office of Educational Research and Improvement, State Accountability Study Group. (1988). *Creating Responsible and Responsive Accountability Systems.* Washington, DC.

U.S. Department of Education, The National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform.* Washington, DC: U.S. Government Printing Office.

Vance, V. S., and P. C. Schlechty. (1982). The Distribution of Academic Ability in the Teaching Force: Policy Implications. *Phi Delta Kappan,* 22–27.

Wenglinsky, H. (1998). *Does It Compute? The Relationship between Educational Technology and Student Achievement in Mathematics.* Princeton, NJ: Educational Testing Service.

Whitener, S. D., K. J. Gruber, H. Lynch, K. Tingos, M. Perona, and S. Fondelier. (1997). *Characteristics of Stayers, Movers, Leavers: Results from the Teachers Follow-up Survey: 1994–95* (NCES 97–450). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Willms, J. D., and F. H. Echols. (1992). Alert and Inert Clients: The Scottish Experience of Parental Choice. *Economics of Education Review, 11:* 339–350.

Wirt, J., S. P. Choy, Y. Bae, J. Sable, A. Gruner, J. Stennett, and M. Perie. (1999). *The Condition of Education,* 1999 U.S. Department of Education. Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement.

Wirt, J., T. Snyder, J. Sable, S. P. Choy, Y. Bae, J. Stennett, A. Gruner, and M. Perie. (1998). *The Condition of Education 1998* (NCES 97–388). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Wirt, J. G., S. Choy, A. Gruner, J. Sable, R. Tobin, Y. Bae, J. Stennett, J. Sexton, S. Watanabe, N. Zill, and J. West. (2000). *The Condition of Education, 2000* (NCES 2000–062). Department of Education. Washington, DC: National Center for Education Statistics.

United States
Department of Education
ED Pubs
8242-B Sandy Court
Jessup, MD 20794-1398

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

Standard Mail (A)