CHAPTER 6

School Effects

Theoretical and Methodological Issues

AAGE B. SØRENSEN
STEPHEN L. MORGAN

INTRODUCTION

Schools are charged with many tasks: the preservation of order through the socialization of children, the maintenance of a productive labor force, the promotion of tolerance, the cultivation of talent, and the prevention of crime and loitering. The satisfactory accomplishment of these objectives, except perhaps the last, is dependent on the capacity of schools to increase individual knowledge, skills, and maturity. Social scientists have many theories about the influence of schools but often little evidence supporting these theories.

The gap between theory and evidence persists for at least three reasons. First, in order to evaluate theories about the influence of schools we need measures of the many possible outcomes of schooling. Unfortunately, we have reasonably good measures for only a few outcomes: cognitive skills and knowledge, economic growth, and labor market success. Second, schools contribute to individual change in interaction with other institutions, most notably the family. The disentanglement of the interacting contributions of schools and these associated institutions is a daunting methodological challenge. Schools have a near monopoly on instruction in some areas, such as trigonometry, and in these areas school effects can be identified.

AAGE B. SØRENSEN AND STEPHEN L. MORGAN • Department of Sociology, Harvard University, Cambridge, Massachusetts 02138

Handbook of the Sociology of Education, edited by Maureen T. Hallinan. Kluwer Academic/Plenum Publishers, New York, 2000.

However, the contribution of schools to psycho-social maturity is clearly very difficult to disentangle from influences of other agents of socialization. Finally, the confounding effects of unobservable individual heterogeneity, especially in genetic endowments, further complicate efforts to isolate the effects of schools.

In this chapter, we discuss theory and evidence on school effects. We address the topic, as it has been addressed by almost all past research in the sociology of education, as a question of how much difference characteristics of schools make for student learning in areas where we have measures of learning outcomes, usually based on performance on standardized tests of achievement. As a result, this chapter, like the sub-discipline of sociology of education, focuses analysis on a subset of the effects of schools noted above. This narrow focus is justified to some extent by the attentions of policymakers and parents, whose decisions—the allocation of funding and residential/school choice, respectively—are influenced by beliefs about the causes of variation among schools in learning outcomes such as mathematics skills and verbal competence.

Research in this tradition is based upon measurement of the impact of differences among schools on learning outcomes. If stable associations of outcomes with differences across schools are established, the conclusion that school effects are strong—a position already held by many policymakers and parents—is supported. However, it is important to recognize that the absence of associations between school differences and learning does not support the conclusion that schools are unimportant. For example, if all schools successfully produce the same level of achievement among students on some outcome, no impact will be inferred from an examination of variation across schools with different characteristics. Nonetheless, the effect of schooling may still be powerful, especially if uniform achievement across schools compensates for inequality in outcomes that would otherwise arise from individual differences in family background and genetic endowments. The possible existence of such a research finding justifies a distinction between the effects of schools (i.e., the effect of variation among schools on outcomes) and the effect of schooling (i.e., the overall effect of the schooling process). When studied at the individual level, the effect of schooling is the major focus of research on the effects of educational attainment. The preoccupation of this strand of research is the disentanglement of the individual effects of family background and genetic ability, not the identification of the influence of different school practices on learning. Interest in the macro-social effect of variation in schooling processes has generated little empirical research, with the exception of research in economics on the effect of education on economic growth that was the inspiration for the early development of human capital theory (e.g., Schultz, 1961). A more recent example of macro-social research on the measurement of the economic effect of schooling has been provided by Angrist and Krueger (1991).

The narrow specification of the research question, however, has the virtue of suggesting research designs which allow for straightforward identification and estimation of the effects of schools. If we could take a sample of schools that differ in organization and/or resources and then either allocate identical students or randomly assign a sample of nonidentical students to various schools, observed variation in learning across these schools would then represent the differential effectiveness of schools. Unfortunately, neither version of this research design is feasible. The identification of identical individuals, even if they existed, is impossible. Furthermore, although experiments abound in educational research, random assignment of students to treatments is rare, perhaps because the design is considered unethical, even if, as is unlikely, parents would allow their children to participate in any such study. There are notable exceptions: the so-called Perry Preschool Project (e.g. Barnett, 1995) on the effect of intensive preschool programs—a large-scale randomized experiment in Tennessee to establish the effect of class size on student performance (discussed in Mosteller, 1995), and the experiments on the expectancy effect by Rosenthal and Jacobson (1968).

The approximation in common use to the ideal experimental design is the statistical estimation of linear models that control for variation, among students, that confounds the estimation of school effects. The use of statistical methods, in particular regression analysis, to control for variation in individual endowments that are relevant for learning has become the standard methodology. Control variables—measured attributes of students and, more recently, assumed functions of unmeasured variables that produce non-random selection of students into schools—are entered alongside the school characteristics of interest in linear-additive models with measures of level of achievement, or gains in achievement, as dependent variables. The only commonly acknowledged weakness of this approach, much emphasized in debates about the magnitude of school effects, is the indeterminate nature of guidelines on which control variables should be included in any model. As a result, the mantra "Many controls are better than few controls" guides most research practice. The avowed purpose of including a maximal set of control variables is to eliminate all possible sources of individual variation that could confound the estimation of a school effect.

Many results have been obtained from research conducted in this manner. However, no consensus exists on the magnitude or source of school effects on learning, especially in relation to the effects of individual endowments. Debates about the relative merits of findings typically focus on methodological issues. While we discuss methodological issues in this chapter, we first examine more fundamental theoretical issues. Theory precedes estimation, for it is theory that determines the type of model that needs to be estimated. Only with a perfectly tailored randomized experiment can estimation proceed without explicit reference to theory, though theory, of course, must guide the experimental design.

The choice of models in regression analysis is usually not seen as a matter of theory by sociologists and educational researchers. Theory, at most, is considered relevant for the choice of control variables to include in a model. Nevertheless, an additive model does represent a theory about how school effects, or effects of any kind, are produced. If a researcher includes a set of school characteristics alongside a set of individual variables, then an implicit assumption is maintained: the effects of school characteristics add to the effects of individual characteristics on learning. In other words, students are provided with additional resources relevant for learning when they enter a school. Students may, for example, be motivated to try harder in schools with certain characteristics. The increased motivation of students may then increase academic achievement, regardless of students' abilities and the opportunities to learn that are provided the students by schools. This may be a reasonable theory about how schools affect learning, but it is not the only mechanism by which learning is achieved. If, instead, learning is jointly dependent on the student's efforts, abilities, and opportunities to learn, then the effect of effort on achievement will depend on the student's opportunities for learning, or how much is taught in a school. This alternative theory suggests that effort and ability will have no effect on learning in schools that teach nothing. The mechanism at the heart of this theory requires a nonlinear specification of the learning process that we later detail. An innocuous sounding assertion, this simple stipulation has strong implications for inferences about school effects in empirical research.

PAST SCHOOL EFFECTS RESEARCH

The school effects literature is large. Any proposal for improving the performance of schools can be considered a theory about school effects, and any study that claims that a school characteristic makes a difference for learning could be said to constitute a contribution to school effects research. In this chapter, we mostly consider large-scale quantitative studies of the

effect of school organization and resources on learning outcomes, and because James S. Coleman's contributions dominate the field, we first survey his seminal research in the sociology of education.

Coleman's Research on School Effects

Coleman began his first study of schools in 1957 when he became an assistant professor at the University of Chicago. A study of adolescent subcultures based on surveys of students in ten high schools in northern Illinois, his first research was published as *The Adolescent Society* (Coleman, 1961). In the preface, Coleman noted two reasons for undertaking the study. Interested in how schools might be made more effective, Coleman (1961) conducted the research because of ". . . a deep concern I have had, since my own high school days, with high schools and with . . . ways to make an adolescent's experience with learning more profitable. . . . " (p. vii). Coleman was also motivated by an interest in different types of status systems and the value systems that they reflect. Coleman's interest in documenting the effects of high schools on learning did not receive much attention in *The Adolescent Society,* possibly because the empirical evidence he gathered about these effects proved more difficult to interpret than he anticipated. Instead, *The Adolescent Society* focuses on social rather than educational processes, cataloguing the social systems created in the ten high schools and in the adolescent subcultures associated with them. Nevertheless, the policy prescriptions Coleman (1961) drew from his analysis were about how to change these subcultures in order to encourage learning and academic achievement.

Coleman's interest in the effects of schools on individual learning became his main preoccupation in educational research for almost four decades. The link between the two concerns motivating *The Adolescent Society* (Coleman,1961)—school effectiveness and schools as social systems—was always clear to Coleman, and his interest in the social systems created in schools reemerges in his last substantial research project in the sociology of education.

Coleman completed three major reports after *The Adolescent Society.* The first Coleman Report (for a while known as the Coleman–Campbell Report) was the massive *Equality of Educational Opportunity*, probably still the largest social science research project ever completed (Coleman et al., 1966). The report's finding of small effects, relative to those of family background, of school resources and facilities on students' achievement created considerable controversy in both academic and policy circles. The second Coleman Report (Coleman, Kelly, & Moon, 1975) was a study of trends in desegregation in American schools. This report did not focus directly on school effects or educational processes but instead focused on the most important policy consequence of the *Equality of Educational Opportunity* report for policy—the use of busing to integrate schools in order to increase the academic achievement of minorities. This second report was widely interpreted to deny the benefits of desegregation documented in the first report and was strongly attacked by many, including many of Coleman's fellow sociologists. The interpretation was incorrect. Coleman did not doubt the relevance of student body composition for schools' educational climates. Rather, he doubted the long-term benefits of busing as a remedy for segregation

The third major piece of educational research Coleman directed, the *High School and Beyond* study, again defined the research agenda by reporting on differential school effectiveness. For this report, however, Coleman and his colleagues focused on differences between public and private schools (published as Coleman, Hoffer, & Kilgore, 1982, and later updated in Coleman & Hoffer, 1987, with additional evidence). The effects of schools, which were

said to have been denied by Coleman in the *Equality of Educational Opportunity* report, were now emphasized. Roles reversed, and Coleman's critics now denied the existence of school effects, at least across the private and public sectors.

The overall impact of Coleman's research in the sociology of education is extraordinary, influencing policy and research as no other sociological enterprise. *Equality of Educational Opportunity* (Coleman et al., 1966) was never replicated, though the data were re-analyzed (see Mosteller & Moynihan, 1972). The public versus private school research inspired other research on school effectiveness, not only in an attempt to reassess Coleman's conclusions, but also to more broadly assess how school resources and organization can affect educational performance and how incentives for improving the performance of schools, such as voucher systems, might promote better school organization and practices. This follow-up research has been primarily conducted by economists and political scientists (Chubb & Moe, 1990; and Hanushek, 1986, 1996). The controversy about the magnitude and source of the effects remains vibrant (Berliner & Biddle, 1997), especially about the incentive systems that are suggested by research. The controversy to date has been dominated by political issues, focusing on the distributional consequences of voucher systems and the ability of traditional interest groups, such as teacher unions and public bureaucracies, to maintain control over the organization and funding of schools.

Most of the controversies surrounding Coleman's school effects research have been based on methodological disputes. As noted above, we emphasize theoretical issues rarely discussed in the relevant literature. Our first task in this chapter is to identify Coleman's implicit theory of school effects.

Coleman's Theory of School Effects

Over the stretch of time in which he was engaged in research on schools, Coleman changed in his ideas about school effectiveness and the link between school effectiveness and social systems. However, there is more continuity in his ideas than is generally recognized. Coleman pursued the same questions, and his alleged reversals of conclusions are much less dramatic than the controversies around these conclusions suggest. The continuity of the questions that guided Coleman's research allows us to identify a theory of school effects on learning.

Coleman may well have been the leader of the discipline of sociology who was most committed to the creation of a science of sociology based on Durkheim's contention that social systems shape and perhaps determine individual action. Coleman's work is not usually identified as Durkheimian, for he also devoted much effort to the development of rational choice theory and methodological individualism where his major concern was the analytic movement from the individual actor level to the macrostructural level (or, conceptually, from microeconomic ideas to Durkheimian perspectives on social structures). However, although building grand theory, his empirical research, including his educational research, made little use of rational choice theory and remained straightforwardly Durkheimian. Although perhaps least apparent in *Equality of Educational Opportunity*[1] (Coleman et al., 1966), Coleman's fascination with the connection between social systems and educational outcomes is most pronounced in *The Adolescent Society* (Coleman, 1961) and in *Public and Private High Schools: The Impact of Communities* (Coleman & Hoffer, 1987).

[1]Coleman deplored this later, characterizing the *Equality of Educational Opportunity* research as not really sociological research (personal communication).

By Durkheimian perspective we mean that Coleman saw social systems as the fundamental determinant of action. Social systems create the outcomes of educational processes by determining individuals' values and motivation. In *The Adolescent Society*, Coleman (1961) argues that the status systems established in peer groups shape students' attention and effort in school. Learning is an individual enterprise with individual level successes measured in grades that indicate relative standing. Success for some students constitutes failure for others. For Coleman, the predominant influence of an adolescent subculture on individual scholastic achievement is negative. Coleman proposed how this might be changed by making scholastic achievement more of a source of collective effort and pride, similar to interscholastic sports. He also, presumably in continuation of these ideas, initiated a research and development effort to create educational games that make social processes support efforts at learning and achievement. However, the evidence is weak for the linkage between social systems and educational outcomes in *The Adolescent Society* (Coleman, 1961). Coleman devotes a chapter to the topic and finds no relationship between the value systems of schools and the performance of students measured as grades obtained relative to the measured I.Q. of students. He presumably took I.Q. to be a measure of a student's intellectual endowment, and then assessed the impact of adolescent culture on performance relative to the potential given by this endowment. Coleman never again used this dependent variable or made a distinction between achievement and ability.

The *Equality of Educational Opportunity* (Coleman et al., 1966) report addressed the problem of schools and their relevance for learning in a very different fashion. As a research report prepared for policymakers, almost nothing is written about the conceptual framework for the analysis. However, on page 36, Coleman writes:

> the question of this report becomes a simple one: How well do the schools of our nation provide such opportunity for minority group children who would otherwise begin adult life with a distinct disadvantage? . . . To answer such a question . . . requires a variety of approaches. Most fundamental, of course, is the question of how well schools reduce the inequity of birth . . . ; that is, what results do schools produce? (Coleman et al., 1966, p. 36).

This is the main statement of the guiding research question, and it also provides an implicit theory of the educational process. The theory is that schools can somehow modify the inequality in educational outcomes created by birth (i.e., created by the family of origin and observable across characteristics such as race, ethnicity, and socioeconomic status). In Coleman's language from *The Adolescent Society* (1961), the question can be formulated as how the social system of the school can modify the social system created by families of origin.

The relationship between family and school was of little concern in *The Adolescent Society* (Coleman, 1961). This relationship was the source of the major finding of *Equality of Educational Opportunity* (Coleman et al., 1966). Another important difference between the two studies was the measure of academic achievement used. While *The Adolescent Society* (1961) saw the impact of the social system of the school on grades relative to a student's I.Q., the outcome in the *Equality of Educational Opportunity* report is academic achievement measured by a standardized test. No independent measure of mental ability was obtained, increasing the burden on the explanatory power of the theory of educational processes being proposed.

As noted, Coleman and colleagues used a largely implicit theory in the *Equality of Educational Opportunity* (1966). Two social systems were compared—one characterizing the family of origin and one created by the school. In practice, the research focus was the measurement of which of these two systems explained the most inequality (or variance) in educational achievement. This question does not have a straightforward answer because variables

characterizing schools are highly correlated with variables characterizing families. Therefore, the result of the decomposition of variance depends on the causal assumptions maintained. The validity of the assumptions became an important controversy surrounding the *Equality of Educational Opportunity* report, as we later discuss. However, the main conclusion is now accepted wisdom: family is much more important for inequality of achievement than the school characteristics that can be purchased with economic resources.

Equality of Educational Opportunity finds that a few school characteristics that reflect the social systems created in schools have some importance for achievement. Consistent with a conjecture of *The Adolescent Society*[2], this finding suggests that those things that can be purchased with money are unimportant, while those reflecting social processes are important. However, the measures of the social system in *Equality of Educational Opportunity*—the average socioeconomic characteristics and racial composition of the student body—are much less sophisticated than those of *The Adolescent Society* (Coleman, 1961). *Equality of Educational Opportunity* presents very elaborate analyses showing that minority students' performance is affected by the composition of the student body—a finding that was to become important for policy and for practice because it was used to justify busing of children to achieve racial balance. In light of the importance of the busing controversies, it is ironic that there was a coding error in the construction of the student body composition variable that may have led Coleman et al. to misinterpret the effect of student body composition (see Smith, 1972).

The main conclusion of *Equality of Educational Opportunity* (Coleman et al., 1966) was based on the decomposition of variance in additive models for the level of academic achievement. This is a particular way to summarize effects in regression analysis. These models are an implicit acceptance of the theory of education that suggests that the influence of families and schools on educational outcomes can be added together. In *The Adolescent Society*, Coleman (1961) saw learning as an outcome that is created by students with different abilities who apply themselves in different degrees to learning. The theory of *Equality of Educational Opportunity* suggests that schools and families create both the abilities and the effort to apply these abilities.

The theory in *Equality of Educational Opportunity* (Coleman et al, 1966) is made even more demanding by the dependent variable adopted. To assess school effects, an outcome measure is needed that can be compared across schools. American schools teach a number of different things to the same age students. In a study as large as *Equality of Educational Opportunity*, the only possible comparable measure is one that is much closer to a measure of student aptitude than to what students have actually learned. Thus, very little of the variation in the achievement measure that is the dependent variable will reflect what students have been taught and thus the outcome of what actually is the main activity in schools. Only to the extent that schools influence student ability and effort, over and above the influence of early socialization and of genes, will schools cause variation in student outcomes measured with these achievement tests.

In the first publication from the *High School and Beyond* study (Coleman et al., 1982), level of achievement was again the main dependent variable. However, gains in achievement were also analyzed, using synthetic gain scores constructed from cross-sectional data. In the second main publication from the study (Coleman & Hoffer 1987), gain scores were analyzed directly. By using gain scores, the demands of the additive theory were reduced. Gains in

[2]In *The Adolescent Society*, Coleman noted with approval an unpublished study from Connecticut showing that per-pupil school expenditure had no relationship to the achievement of students relative to their I.Q.

achievement are outcomes schools presumably can influence, assuming achievement is measured so it reflects what can be learned in schools. A measure comparable across all schools is still needed, and the tests used in *High School and Beyond* (Coleman et al., 1982) seem to measure aptitude, reducing the direct connection to what may be taught in schools. However, Heyns and Hilton (1982) argued that the tests do measure something schools can influence.

The mechanisms that produced the dependent variable again did not receive much attention in the *High School and Beyond* (Coleman et al., 1982) research. Coleman and Hoffer (1987) showed that private and public schools differ in many ways: in course offerings, in organizational differentiation, in parental involvement, in discipline, in homework, size, and so on. However, these characteristics were never directly related to achievement outcomes in a manner that suggests how they affect achievement. Instead, theoretical attention was devoted to the specification of how the social systems created in schools can become effective agents for creating student behavior that is conducive to learning. Thus, the main focus is the same as in *The Adolescent Society* (Coleman, 1961). In *Private and Public High Schools* (Coleman & Hoffer, 1987), Coleman changed the emphasis of what was important about school social systems. The direct emphasis on peer-group cultures (in *The Adolescent Society*) and on student body composition (in *Equality of Educational Opportunity*) was replaced by attention to the attributes of the school social systems that could make them extensions of the family, in order to create a mutually reinforcing community.

Coleman and Hoffer (1987) noted that some schools simply extend the values of the family; nonreligious private schools and some public schools of choice attract students on the basis of their own and their parents' values. In contrast, some religious private schools, Catholic schools in particular, not only share values, but are a part of a functional community that reinforces these values. Religious schools are high in social capital, a concept that became important for Coleman and for others. These schools are easy to manage and their pupils are easy to teach. The social structures that envelop students generate beneficial effects on students' learning.

The basic form of the new theory is the same. The network structures that create social capital add together family and school effects to create powerful social influences that promote educational outcomes. Students immersed in social systems rich in social capital gain in academic achievement presumably because they work harder and with greater aptitude in schools rich in social capital. The theory is perhaps a bit less ambitious than the theory implicit in *Equality of Educational Opportunity* (Coleman et al., 1966). However, it is much more ambitious than the theory of *The Adolescent Society* (Coleman, 1961), for there schools are only asked to affect the grades that students obtain relative to their ability.

In *Public and Private Schools* (Coleman & Hoffer, 1987), the criteria for testing the theory also changed. Coleman no longer gauged the relative explanatory power of family and school. Instead, he assessed the significance of the effect of social capital. Unfortunately, the used data had no direct measure of social capital. As a result, Coleman and his associates could not directly test their social capital hypothesis and instead relied on an interpretation of the association between school sector and achievement gains. When the assertion that Catholic schools have the highest social capital is coupled with the empirical finding that Catholic school students score slightly higher on standardized tests of achievement, the theory of social capital seems to have some support. Methodological attention therefore focused on establishing the significance of sector effects, net of the selection of students with different abilities into different school sectors. Because no direct measure of ability is available in the *High School and Beyond* data, Coleman and his associates use measures of family background as proxy variables for ability (and perhaps other confounding variables).

In Coleman & Hoffer (1987), the family background variables are used in an additive model with gains in achievement as the dependent variable. The model is then estimated separately for each school type. This implies an interaction between school type and the influence of family background (and achievement at the sophomore level) on gains between sophomore and senior year. The conjecture that high social capital supports family and school structures might suggest to some that the interaction should be such that the family background variables are more important in schools with strong ties between the family and the school system. However, Coleman & Hoffer (1987) evidently conceived of the matter in just the opposite manner. Schools with high social capital, such as Catholic schools, should reduce the influence of family background variables, he argued. In the language of *Equality of Educational Opportunity* (Coleman et al., 1966), the variance explained by family variables should be lower in Catholic schools. A smaller effect of family socioeconomic status on achievement gains is indeed found for Catholic schools than is found for public schools, supporting the conclusion that parochial schools are common schools that reduce the impact of social origin.

The theory of schooling implicit in Coleman & Hoffer (1987) project is more complex than in the *Equality of Educational Opportunity* report. The earlier theory suggested that academic achievement was produced by two sets of forces—one set formed by the family and another by the school—and the salient question is the relative importance of the two. The new theory suggests that the degree of integration of the community social system, formed by school and family structures, determines the degree to which the comprehensive community social system produces learning and reduces the deficits some children bring to schools. The implicit model is one in which high social capital adds to the academic achievement of everyone and reduces the importance of family background. In other words, some school and community characteristics interact with family background.

The argument for the importance of an interaction effect between school characteristics and family background seems indeed the proper specification of the statement of the *Equality of Educational Opportunity* report: "Most fundamental, of course, is the question of how well schools reduce the inequity of birth. . . . " (Coleman et al., 1966, p. 36). This question could not be answered in *Equality of Educational Opportunity* because Coleman never interacted school characteristics with family background characteristics.

The specific interaction between school characteristics and family background, implicitly proposed by Coleman, is compensatory in nature. Good schools reduce the inequity of birth because the social systems created by good schools reduce the importance of family background. However, this is not the only possible direction of the effect of the interaction between school characteristics and family background characteristics, as we will demonstrate below.

The Adolescent Society (Coleman, 1961) research considered school effectiveness as a question of how schools might affect the relationship between I.Q. and achievement. If we consider I.Q. a measure of family background, then the question is exactly the same as the one addressed by *Public and Private Schools* (Coleman & Hoffer, 1987) and formulated, but not addressed, in *Equality of Educational Opportunity* (Coleman et al., 1966). There is indeed consistency in Coleman's theory of how schools affect educational processes.

Conceptions of School Effects in Other Research

Coleman's school effects research defined the basic questions addressed by most other school effects research. Hanushek (1986) reviewed 147 studies of the effect of school characteristics

on learning and found no systematic pattern of the effect of class size, teacher/pupil ratio, teacher education and experience, teacher salary, and per pupil expenditure on achievement. Most of these studies used administratively collected data from schools and school districts. The data showed a positive correlation between school expenditures and achievement, but revealed that the strength of the relationship disappears when family background is controlled.

Hanushek (1986) also reviewed several classroom-level studies that showed clear differences in teacher effectiveness (e.g., Hanushek, 1971; Murnane & Phillips, 1981). Some teachers do significantly better than other teachers in improving students' achievement test scores, controlling for background characteristics of these students. However, these total effects of teachers are not related to easily observable characteristics of teachers, such as their education and experience. Teacher effort and ability at teaching does appear to matter, but the relevant characteristics are not identified in the measures used in large scale research.

Other varieties of research have been pursued over the last few decades. A number of studies that investigated learning outcomes in one or a few schools claimed to provide evidence of what makes a school effective (e.g., Goodlad, 1984; Lightfoot, 1983; Powell, Farrar and Cohen, 1985, Sizer, 1984). These studies tended to be cross-sectional, based on small samples, and qualitative. They therefore fell outside of the review we are undertaking. Bryk, Lee & Holland (1993) focused on the common-school conception of Catholic schools, providing both qualitative and illustrative quantitative evidence.

Much attention has been focused on the average performance of U.S. schools, both over time and in comparison with the schools of other nations (for an overview of this evidence see Hanushek, 1994a). The former is investigated by comparing SAT scores and other average achievement test scores over time for the nation as a whole. The trend is a marked decline in SAT scores, but this partly reflects changes in the composition of test takers. Comparisons of achievement test scores given to representative samples show very little change. Nonetheless, the gap between Whites and minority groups has narrowed over the last three decades (Jenck & Phillips, 1998). Comparisons with the schools of other nations show that US students perform worse than students from many other nations. An important characteristic of U.S. schools revealed in these comparisons is that there is a great deal of variation in performance and in student body composition among states (see National Center for Educational Statistics, 1993). There is also a great deal of variation among U.S. schools in what schools teach and require in mathematics of their students. This variation presumably accounts for a major part of the difference between the United States and nations with more uniform standards and curricula. When course taking is taken into account, U.S. students taking advanced courses do as well, if not better than the average Japanese student (Westbury, 1992). If all students learned from the same textbooks, the U.S. deficit might also decrease. Schools decide what textbooks to use and can require students to take certain courses. Thus, the practices of American schools do matter and can be altered.

Throughout the long period of stagnation in performance, spending on schools has increased dramatically. Thus, schools deliver less per dollar than they did previously. Much of the debate about this lamentable state of affairs has focused on providing incentives for schools to do better. A large number of innovations—charter schools, magnet schools, school choice, performance contracting, and voucher systems—have been proposed, but research evaluating these innovations is (so far) inconclusive. Whatever insight this research provides about the effectiveness of the incentive systems and innovations, it provides very little insight into what specifically schools can do to improve their own incentives to educate students. A notable exception is the research conducted by Chubb and Moe (1990).

Chubb and Moe (1990) suggested an elaborate theory of how variation in school effec-

tiveness reflects the institutional context of schools, especially the institutions of direct democratic control by which schools have traditionally been governed. This system, and the bureaucracies it creates, constrains the shape of school organization that is important for learning outcomes. In contrast to the effective schools literature, which makes similar claims with qualitative evidence, Chubb and Moe (1990) offered quantitative evidence using the same *High School and Beyond* data analyzed by Coleman (1961). In contrast to the private versus public school research conducted by Coleman and associates, which hypothesized that school organization (shaped by social capital) accounts for the difference between private and public schools, Chubb and Moe introduced measures of what they consider to be the relevant aspects of school organization, believed to be related to gains in achievement. They used the same basic model as Coleman and Hoffer (1987)—a linear model for gains in achievement where measures of school characteristics are introduced alongside measures of family background variables and initial level of achievement.

Chubb and Moe delineated four sets of organizational characteristics of schools: personnel, goals, leadership, and practice. They constructed measures of each set of characteristics by comparing high achieving and low achieving schools and then build a composite index of effective school organization. Finally, they showed that this effective school measure is a strong predictor of gains in achievement. According to their results, the predicted achievement gain, resulting from moving from the lowest to the highest quartile on effective school organization, is about equal to the gain that would be achieved by moving from the lowest to highest quartile of family background. As a result, they concluded that the effect of good school organization is comparable in size to the influence of each student's family background. They also found no effect of racial composition. The findings of Chubb and Moe, if true, are a substantial revision of the main conclusions of the *Equality of Educational Opportunity* (Coleman et al., 1966) report. However, it is important to note that these results on 10th-grade to 12th-grade achievement gain are obtained net of academic achievement in Grade 10, which Chubb and Moe (1990) interpreted as a measure of ability. A substantial effect of family background presumably is mediated by 10th-grade scores. Also, Coleman and associates (1966) used variance explained, rather than estimates of regression coefficients, as the basis for their conclusions.

Chubb and Moe (1990) proceeded to show how effective school organization depends on the institutional setting, as this was their main argument. However, with only one exception, they were unable to cast much light on what schools with effective organizations actually do to produce more learning. They found no effect of the amount of homework required, graduation requirements, administrative routines in the classroom, or discipline. They did find a substantial effect of the percentage of students in a school who are educated in the academic track. Schools in which students take more demanding courses produce more learning. A similar difference in track organization also accounts for most, if not all, of the difference between public and private schools (Alexander & Pallas, 1983; Coleman & Hoffer, 1987). Thus, the only direct effect established between effective school organization and learning is that effective schools have more demanding academic programs. Something other than the instructional program may explain why schools with high goals, functional communities, strong leaders, dedicated teachers, involved parents, and professional teachers produce more learning. However, neither Coleman and Hoffer (1987), nor Chubb and Moe (1990), demonstrated the direct relevance of these organizational variables for learning. The importance of teachers, academic programs, and curricula suggest a simple theory about how important school effects come about. Schools that teach more material produce more learning. In the next section, we further develop this theory.

TOWARD A THEORY OF SCHOOL EFFECTS

Conclusions from school effects research have shifted from the claim that schools make little difference for learning to the alternative claim that organizational features of schools make a lot of difference. Underneath the most recent claims is the relatively uncontroversial finding that the main effects of schools are effects of their academic programs. Differences in teaching programs are also what appears to account for the results of the international comparisons of the performance of U.S. schools with the schools of other nations. In short, schools make a difference by teaching students.

This simple conclusion suggests that research on school effects should pay more attention to variables measuring teaching effort and curricula—variables that measure the opportunities for learning provided by schools. When these variables are added to the regression equations typically estimated in school effects research, opportunities for learning should further explain achievement. However, simply adding variables providing measures of opportunities for learning to the standard additive models is not necessarily the correct specification. The effects of opportunities for learning may not simply add to the effects of student endowments. Rather, it is more likely that the effects of opportunities for learning interact with the effects of student endowments.

Suppose, to take an extreme example, that we have a school in which no English is taught. All instruction is in Hungarian and about Hungarian topics. Furthermore, parents who send their children to this school speak Hungarian and nothing but Hungarian. If given the standard English verbal ability test analyzed in educational research, students from this school would all achieve the same score. Students would not have been taught anything relevant for performance on the English test. No effect of family background on learning outcomes in this school would be observed, for no learning relevant for the test has taken place.

Suppose, in contrast, that there is a school with an extremely rich curriculum in mathematics. Students can progress from very simple algebra to very advanced mathematical topics. Every progression in this curriculum depends on how well a student has learned the previous step, and the speed of learning of each step depends to some degree on the student's background (because his mental ability is correlated with his family's socioeconomic status). The cumulative nature of the material, coupled with the relevance of ability for the completion of each stage in the curriculum, implies both that learning increases with the number of stages in the curriculum and that the correlation between mental ability and the amount learned increases with the number of stages in the curriculum. The assumption of a correlation between family background and mental ability ensures that, other things equal, the more instructional material that is covered in a period of instruction, the higher the observed effect of family background on academic achievement.

These simple ideas suggest that there are two components to the educational process in schools: what students are taught and how much they learn of what they are taught. Schools, classrooms, and other instructional groups differ in what and how much is taught. Students will differ, by ability and effort, in how much they learn. Student ability and effort can change over time as a result of learning and as a result of motivational processes associated with the social systems to which they belong. Schools can influence how much a student will learn by how much they try to teach and by changing student effort.

A simple formal representation, originally proposed in Sørensen and Hallinan (1977), expresses these ideas. Denote by $y_i(t)$ the amount a student, i, has learned by time t of material in an instructional period, or a course, where instruction began at Time 0. Let $v_c(t)$ be a measure of how much material has been presented by time t in classroom c. The amount, $y_i(t)$,

students will have learned of this material in the period from Time 0 to Time t depends on students' ability and effort. Denote jointly ability and effort expended by s_i, and assume, for now, that s_i remains constant throughout the learning process. Ability and effort of course vary across individuals.

Learning is related to teaching by the simple differential equation:

$$\frac{dy_i(t)}{dt} = s_i dv_c(t) \tag{1}$$

The increase in achievement, $dy_i(t)$, for each increment of time, dt, is a function of how much has been taught, or $dv_c(t)$, and on the students's ability and effort, s_i. Using simple methods of integration, the solution is

$$y_i(t) - y_i(0) = s_i[v_c(t) - v_c(0)], \tag{2}$$

which gives the amount learned by time t as a function of how much has been taught in the period and of the student's ability and effort. The simple formulation of Equations (1) and (2) provides one elementary but fundamental insight. Opportunities for learning interact with the effort and ability of students in producing learning. If nothing is taught, nothing is learned. Individuals have to learn from something, perhaps their own experiences, perhaps their textbooks, perhaps their teachers. If they have nothing to learn from, they will learn nothing. Moreover, ability and effort determine how much will be learned in a period of time. If two students are exposed to the same material, the student with the highest ability and/or effort will learn the most. Ability and effort are here conceived of as individual characteristics, varying among students, but treated as constant over the period of time studied. Because there should be no ambiguity about the usage, the subscript *i* is dropped in the sequel. The amount of material taught by time t varies across classrooms because of differences in curricula and in teacher effort. In most of the discussion that follows, $v_c(t)$, is treated as a school characteristic and within-school differences between classrooms are ignored. In empirical work, it may be possible to estimate the variation in v(t) across both schools and classrooms depending on data availability. However, the presentation here is theoretical and we therefore also drop the subscript c in the remainder of this chapter.

The implications of the interaction between ability and effort, on the one hand, and opportunities for learning, on the other, may be seen by introducing variables that are correlated with the ability and effort of students. Assume that we have measures of a student's endowments, such as family background variables and other variables (perhaps school characteristics) that influence the efforts of students. Assume further that we are uninterested in how these variables actually produce ability and effort, so that a linear formulation is adequate. Thus, we specify s as $s = c_0 + c_1x_1 + c_2x_2 \ldots c_nx_n$, and thus re-express Equation (2) as

$$y(t) - y(0) = [v(t) - v(0)](c_0 + c_1x_1 + c_2x_2 \ldots c_nx_n) \tag{3}$$

The regression of $y(t)$–$y(0)$, or gains in achievement over a period, on variables that measure not only ability and effort but also school characteristics is the common practice in school effects research. Equation (3) shows that if the simple mechanism for learning proposed here is correct, then the coefficients of the x variables will depend on $v(t)$–$v(0)$, the amount of instructional material presented in the period over which the gain in achievement is measured. The larger $v(t)$–$v(0)$ is, the larger the estimates of these coefficients will be. Schools

that cover extensive material in their instruction provide many opportunities for learning and will create more growth in achievement. If opportunities for learning are not measured and included in the model, the estimated coefficients of the x variables will reflect these unmeasured and omitted variables. More opportunities for learning will create larger observed effects of the variables that measure the individual endowments that determine ability and effort.

There are several objections that may be raised against the formulation proposed by Equation (3). First, it assumes that effort and ability are constant over the period studied. This is not a reasonable assumption if a student's ability to learn new material depends on what is already known and a student's effort depends on past success. These are both reasonable mechanisms. However, for short periods of time, such as the 2-year period covered by the panel data commonly analyzed in school effects research, the assumptions of stable effort and ability may be reasonable.

A much more serious problem is that direct measures of opportunities for learning are usually not available. Data available for research on school effects usually do not measure what and how much schools teach, though an instructive exception is provided by Barr and Dreeben (1983). This makes it impossible to estimate Equation (3) directly. The solution to this problem adopted in Sørensen and Hallinan (1977) is to assume a particular dependency of $v(t)$ on time. Assume that the amount of new material presented, $dv(t)$, declines over time in proportion to what has already been taught, or $dv(t)/dt = b\, v(t)$, with the constraint that $b < 0$. This expression is again a simple differential equation. When solved, see Sørensen and Hallinan (1977) for details, one obtains the function

$$v(t) = \frac{1}{b}\,(e^{bt} - 1). \tag{4}$$

By letting $t \rightarrow \infty$ as less and less new material is presented, we can show that the total amount of material presented is $-1/b$, so that b directly determines the opportunities for learning. The smaller b is in absolute magnitude, the more opportunities there are. The formulation (4) assumes that $v(0) = 0$, but the specification implies a simple linear differential equation for learning, $dy(t)/dt = s + by(t)$, that can be solved for any period of time. The solution is

$$y(t_2) = c_0^* + b^* y(t_1) + c_1^* x_1 + c_2^* x_2 \ldots . c_n^* x \tag{5}$$

where $b^* = e^{b\Delta t}$, and each of the c^* coefficients have the form, $c^* = c/b(e^{b\Delta t}\,\Delta 1)$. Here $\Delta t = t_2 - t_1$, the period of observation (e.g., 10th to 12th grade). From estimates of b^*, one can obtain b as in $b^*/\Delta t$, and using this estimate, one can also generate the parameters c from the c^*s. Note that b^* is a function of b and time. Estimates of b^* are between zero and 1, since b is assumed to be less than zero, For given time periods, a larger b^* suggests more opportunities.

Equation (5) is the lagged regression model estimated by Coleman and Hoffer (1987), Chubb and Moe (1990), and many others. However, the proposed derivation, and the resulting parametrization, allows for the estimation of the fundamental forces governing learning: the opportunities for learning characterizing schools and the individual endowments of students, perhaps augmented by their schools. Such estimates were presented for various school characteristics in Sørensen and Hallinan (1977), for ability groups in Sørensen and Hallinan (1986), for academic tracks in Sørensen (1987), and for public versus Catholic schools in Sørensen (1996). All of these applications provided results consistent with the ideas suggested above. The models were estimated individual level models. Multilevel models would have been more

appropriate because the b^*s are school (track or ability) characteristics. This would have provided better estimates, but the use of multilevel analysis would not change the conceptual points emphasized. In the analysis of the Catholic versus public school differences, estimates for b in different subject areas suggest that Catholic Schools do indeed provide more opportunities for learning, in addition to whatever effects the functional communities have on the efforts of students (Sørensen, 1996). This appears to be mainly a result of more students being allocated to academic tracks in Catholic schools, as noted above.

The conception of how school effects are produced proposed here is very different from what Coleman proposed in the additive models used in *Equality of Educational Opportunity* (Coleman et al, 1966) and subsequent school effects research. The additive models used assumed that schools can shape the minds and motivations of students in order to overcome the influence of family background. Additive models pay no attention to the interaction between student endowments and opportunities for learning, or the process of education.

In response to an early version of the model proposed, Hauser (1978) argued that it is unreasonable to assume that school effects are produced primarily by opportunities for learning. He suggested that there may be considerably more learning going on in a boot camp where relative achievement levels are in flux than in a perfunctorily led high school class where relative learning differences persist. Evidently, Hauser conceived of learning as change in students' ability and effort only. Furthermore, Hauser seemed only to allow school effects that change students relative to each other. This is an unreasonably restrictive definition of school effects: schools only matter if they make the bright dull and the dull brighter, as can be detected by the variance partitioning produced by simple cross-sectional models.[3]

Hauser (1978), Coleman and associates (1966), Chubb and Moe (1990), and others conceived of schooling as a process whereby schools somehow add to, or subtract from, the intellectual resources of students. This is an attractive scenario, for it implies that schools should be able to produce gifted students and more equal educational outcomes. At least all students attending the same schools should come out pretty much alike, and all students in America should become equal if schools are equally equipped with the things that mold young minds. This, in turn, would mean that schools can achieve equality of results, as well as equality of educational opportunity, if those who are in charge of schools think that this is an important goal.

This conception of the educational process and of the role of schools is an important one in American culture. The comprehensive system of secondary education, or the "common school" (Cremin, 1951), is a unique American institution designed to achieve a basic equality of educational outcomes. The institution has been imitated in a number of European countries, as a replacement for, or an alternative to, the very selective and highly differentiated European school systems, designed originally to achieve the opposite of the American schools: maximum feasible inequality of educational outcomes (see, for example, Kerckhoff (1993), who presented an exemplary illustration of the contrast in an empirical analysis of schooling processes in the U.K., emphasizing the long-term consequences of processes creating inequality).

The American goal of the common school is not an easy one to achieve, at least not if it is taken literally. Presumably, schools would add to the ability and effort of children primarily by teaching and thus by creating opportunities for learning. If equal outcomes are desired, these opportunities should be allocated so that the least able receive the most opportunities and the most able the fewest. This is not usually what occurs in schools. One main reason may

[3]A number of implications for policy of the debate between Hauser and Sørensen about the nature of school effects are developed by Hoaglin et al. (1982).

be that such an allocation of opportunities is in direct contradiction to another popularly accepted educational goal; the goal of providing each student with the opportunity to achieve to the maximum of his or her potential.

Providing students with more opportunities for learning produces more inequality in academic achievement. This is easily seen from the conception of learning and teaching proposed here. Opportunities for learning determine the parameters of Equation (5) in such a manner that the variance of y(t) in the long run will become $s^2_{y(max)} = (-1/b)^2\sigma^2_s$. In this case, y(max) is the eventual academic achievement, and σ^2_s is the variance in student intellectual resources. Clearly, as more opportunities are provided, more variance or inequality in academic achievement will be created, for given inputs. From Equation (5), it can also easily be seen that the variance in achievement will increase over time until it reaches the value $\sigma^2_{y(max)}$. There is empirical support for this prediction about the increase in inequality in achievement in the form of a phenomenon called fan spread. For example, Willms and Jacobsen (1990) reported that while students tend to maintain their initial position in the distribution of mathematics achievement from grades 3 to 7, there is an increase in the variance in the scores in later grades that is consistent with the suggsted predictions suggested.

In fact, the goal of providing each student with the opportunities to achieve the maximum of his or her potential has stronger implications. Such a goal implies that those with the greatest intellectual resources should have the most favorable opportunities for learning. For this reason, ability grouping and similar arrangements are often adopted in the lower grades in tandem with elaborate curriculum differentiation in the higher grades. Schools that provide the most opportunities to the most able students are also schools that maximize differences among students in academic achievement. The goal of maximizing each student's academic achievement, therefore, implies the maximization of inequality of educational outcomes given individual endowments.

If we consider good schools to be schools that try to teach a lot, and if the model proposed previously is true, then good schools increase the effect of family background. This is, of course, the exact opposite of Coleman's prediction (Coleman et al., 1966). His conception of a good school does not emphasize the amount of teaching taking place, but the strength of functional communities that may be created in schools. These characteristics of good schools reduce the effect of family background, he argued. In Catholic schools the effect of family background on learning should be lower than in public schools. Coleman, Hoffer, and Kilgore (1982) and Coleman and Hoffer (1987) showed that there seems to be an advantage for minority students and to students with disadvantaged family backgrounds who attend Catholic schools. There are two explanations for this result. Coleman's would be that it reflects the higher effort by all students produced by functional communities. However, the pattern could also result from selection into Catholic schools by highly motivated parents of disadvantaged youth, as argued by Goldberger and Cain (1982), so that the allegedly disadvantaged students are, in fact, nontypical as shown by their parents' school choices. Coleman did not address these selection issues related to his common school idea (see Heckman & Neal, 1996, for a discussion).

If the smaller effect of family background in Catholic schools is not due to selection, then the result suggests that the theory about the effect of opportunities is wrong, or that Catholic schools provide fewer opportunities for learning, or that the functional communities overcome the increased differentiation caused by better opportunities for learning in Catholic schools. It can be shown, using inclusion in the academic track as an indicator of opportunities for learning, that there indeed is a stronger effect of family background on achievement for students in the academic tracks compared to other students (Sørensen, 1987), suggesting that the interaction effect between background and opportunities for learning does indeed exist. It

is also possible to show that Catholic schools provide more opportunities for learning than public schools (Sørensen, 1996). Thus, it would seem that selection bias is indeed important, or that functional communities change students substantially, as suggested by Coleman's Durkheimian educational theory. In other words, if there is no selection bias, then Coleman & Hoffer's (1987) common school scenario implies that functional communities make students so much alike that the increased opportunities for differentiation provided by the richer teaching in Catholic schools do not result in more inequality in these schools, but in less.

METHODOLOGICAL ISSUES

No other sociological research has attracted as much attention and controversy as the school effects research conducted by Coleman and his associates. Almost all of the debate surrounding these studies focused on methodological issues of study design, measurement, sampling, and statistical analysis techniques. The magnitude of reported effects was disputed. In the first Coleman report, these effects were minimal. In the public/private school research, they were substantial. Much of the debate over these effects focused on whether regression analysis adequately controls for the individual endowments that influence learning outcomes. This section briefly surveys these issues, as well as an unresolved statistical problem of the lagged regression model that has become the common model for the estimation of school effects.

The Magnitude of School Effects

The well-known conclusion of *Equality of Educational Opportunity* (Coleman et al., 1966) is that school resources—instructional and economic—make very little difference for the academic achievement of students. This conclusion was drawn from a comparison of the amount of variance explained by the set of variables measuring school characteristics and the amount of variance explained by individual endowments. The choice of this methodology clearly reflects the statistical and computational capacities that were available at the time of the study. The use of regression analysis by social scientists on large datasets was still in its infancy in the early 1960s. Coleman needed a way to express the effect of a large group of variables simultaneously, and the partitioning of explained variance seemed a straightforward and feasible approach given the computational resources that were available. Using this measure of effect, *Equality of Educational Opportunity* (Table 3.24.1) reported that district per pupil expenditure accounts for a mere .09% of the variance in academic achievement for Black students in the North and .29% for White students in the North.

There are two problems with this approach. One received much attention in the debate about the *Equality of Educational Opportunity* report—the causal ordering of the various groups of variables. Because schools with good resources tend to have students with high family background, groups of variables summarizing each are therefore highly correlated and share substantial amounts of variance. The total amount of variance explained by a group of variables is therefore a poor measure of the importance of the group of variables. The order in which the variables are entered becomes important. Hanushek and Kain (1972, p. 127) showed that school characteristics account for 8.08% of the variance in achievement, if this group of variables is entered first. In contrast, the family background variables account for 7.77% of the variance when they are entered first. When entered last, school resources and instructional facilities account for an additional 2.12 % of the variance.

Coleman and his collaborators on *Equality of Educational Opportunity* (1966) defended their assumptions of causal order by pointing out that family background is causally prior to school characteristics. However, the opposite may be true if students choose schools freely and their choices depend on the facilities of the schools, as noted by Hanushek and Kain (1972). The assumption of Coleman and his colleagues was embedded in the standard interpretation of the results of the study, and for a group of neighborhood schools, the assumption may be reasonable.

However, the interpretation remains ambiguous. The amount of variance explained by schools would be 100% if schools caused all the variation in educational outcomes so that schools produce equal educational outcomes for all their students, regardless of student body composition. However, the outcome of explaining 100% of the variance could also result from schools having no effect at all. This would occur if schools were completely homogeneous in terms of student bodies so that students are all alike within schools and different between schools. In the latter case, even if schools have no effect on academic achievement, school characteristics would perfectly predict the achievement differences caused by individual endowments. If there was no variation within schools in individual endowments, all of the effects of these endowments on achievement would mistakenly be attributed to schools in the standard regression analysis.

There is no way to differentiate these two interpretations without recourse to theory, a randomized experiment, or an unverifiable assumption that we have measured all determinants of achievements and the standard linear additive specification is the true model. Standard school effects research offers none of these alternative guaranties.

The use of amount of variance explained as an effect measure has other problems. Variance explained depends on the variances of the variables. Anyone browsing through the *Equality of Educational Opportunity* report (1966) will find a bewildering variation in the R^2s and contributions to R^2s among the various regional, grade, and racial groups that are largely uninterpretable. These variations clearly reflect variation across groups in population variances, errors of measurement, and the like. None of this variation says anything about the effectiveness of schools. Cain and Watts (1970) pointed this out and argued for converting the variables into some dollar metric and the use of (metric) regression coefficients when evaluating their policy implications. Coleman (1972) converted the variance-explained measures to standardized regression coefficients, for a long time popular measures of relative importance. These measures suffered from the same dependence on population variances as the variance measures.

More recently, school effects research has relied on regression coefficients to assess the magnitude of effects. In the private/public schools research, almost all attention has been focused on the statistical significance of coefficients, in particular those for school sector. Statistical significance is important for our willingness to believe in a result. It is not a guide to the magnitude of the effect because statistical significance also depends on population variances, on sample variances, and on sample size. There nevertheless has been a tendency to confound statistical significance with substantive importance.

Coleman and Hoffer (1987) converted the regression coefficient estimates into grade equivalents and claimed that average public school students would gain from 1 (in reading, writing, and mathematics) to .02 (in science) grade equivalents if they were instead enrolled in Catholic schools. This seems to mean that in important subjects, Catholic school students gain an extra year of schooling over public school students. Over the 2-year period covered by these data, this is a large effect of Catholic schooling, perhaps too large to be believable.

The private/public school research by Coleman and his associates makes no attempt to

compare school effects to family background effects, the issue that so much occupied them in the *Equality of Educational Opportunity* report (1966). School effects are in a different metric than the family background variables, and direct comparison is impossible based on raw regression coefficients. Chubb and Moe (1990), nevertheless, argued that their school organization variable is important because if one moves from the bottom to the top quartile in school organization the gain is even greater than when moving from the bottom to the top quartile in family background. This is a comparison no different from the comparison of variance explained. Unless one believes that effort expended by moving around quartiles is a proper metric, the effects remain incomparable.

Effects of dollars invested in a school might be an appropriate metric because its effect can be compared to investments in other activities. Hanushek (1986), as noted previously, found no systematic effect of school expenditures on learning. His meta-analysis was of patterns of significance and not of the magnitude of effects. Hedges, Laine, and Greenwald (1994) reanalyzed the same set of studies with a different methodology and found a different pattern of effects in terms of statistical significance, suggesting more support for an effect of school expenditures. They also attempted to estimate the magnitude of the effect of dollars invested in schools and found that a 10% increase in real resources produced a 70% increase in a standard deviation of achievement. Hanushek (1994b), in his response, pointed out that the expenditure on schools increased by 100% over 25 years, thus suggesting an increase of seven standard deviations in achievement! In fact, performance has at best remained constant over this 25-year period.

Parents believe that schools make an important difference, important enough to justify additional expenses for private schooling and housing choices to allow their children to go to what they consider to be good public schools. The result is a confounding of school characteristics and individual endowments that has occupied school effects research ever since the publication of *Equality of Educational Opportunity* (Coleman et al., 1966). It is probably meaningless to say which is more important. It is not meaningless to compare expenditures on schools to expenditures on social programs providing a higher standard of living for the poor and more income security for them. These comparisons remain to be done.

Selection Issues

The ideal design for the establishment of school effects would seem to be the randomized experiment noted in the introduction. Short of the ideal design, researchers rely on the use of control variables in regression analysis, usually measures of individual endowments. If these controls are inadequate or are improperly specified, estimates of school effects are biased. This issue has occupied much of the methodological criticism of school effects research and started with the debate over the *Equality of Educational Opportunity* report (Coleman et al., 1966) and its assumption about the causal order of family background variables and school characteristics. Again, the issue dominated the attention of the debate about the research on public and private schools, but by then a new language had been developed by statisticians and by econometricians to explain the sources of bias. The main criticism of the research was that the controls used by Coleman were inadequate, resulting in sample selection bias (see, for example, Goldberger & Cain, 1982; Murnane, Newstead, & Olsen,1985).

Coleman, Hoffer, and Kilgore (1982) used 17 family background variables as controls. However, this did not seem adequate to the critics who pointed out that unmeasured endowment variables might still bias upward the effect of private schools. A set of sophisticated

statistical techniques to handle sample selection bias was being developed at the time Coleman's critics attacked the public and private school research (see Winship & Mare, 1992, for a review of these techniques). There are two issues involved—model misspecification and unobservable causes of self-selection. If the nonrandom allocation of students to public schools and to private schools is not adequately controlled by entering the correct specification of all relevant family background variables and other observable characteristics of students (i.e., parents' education is included as a linear additive effect when a nonlinear specification is more appropriate), then bias in the school effects coefficients can result. The second source of bias in the estimates of school effects can arise from the nonrandom allocation of students to school sectors on the basis of an unobserved determinant (i.e., parents are more likely to pay for what they consider a high-quality private education if they rightly judge their children to have an unusually strong genetic endowment of cognitive ability). The first source of bias can be handled with a technique developed by Rosenbaum and Rubin (1983) to estimate the sector effects separately for groups of students having the same propensity to attend private schools, although constructing (and defending the chosen construction of) the propensity score can be quite difficult. Coleman and Hoffer (1987) attempted this sort of correction and reported that it did not change their conclusions. The second source of bias can be tackled by a two-stage estimation procedure, associated with the labor economics research of Heckman (1979), that supplies an estimate of the error term with which the school effects variables are purportedly correlated. Applied by Noell (1982) to the private/public sector effects estimate, no significant impact of unmeasured variables on the sector effect was detected.

Finally, control variables can also magnify selection bias, as often happens when researchers do not use an explicit theory to model the mechanisms that produce school effects. We proposed previously that schools may affect learning in two ways—by providing opportunities for learning and by changing the ability and effort of students. These latter atmospheric effects of schools presumably also affect parents. With the desire to control for all relevant characteristics of family background, researchers often include family-level variables that might be partly determined by schools. Such variables include items in the home—number of books, dictionaries, pocket calculators—all items used in the research on private and public schools. Similarly, parents presumably react to the ability and effort of students by adjusting their aspirations and expectations to the school performance of their children. These variables are commonly used to measure family background. Finally, even the standard family background variables, parents' education and socioeconomic status, may change meaning as a result of the operation of schools. These variables are indicators of cultural aspects of the home background, and if schools change students, schools also change what these variables measure.

Statistical Issues

The estimation, in early schools effects research, of linear models for the level of academic achievement using ordinary least squares presents no special statistical problems. However, in almost all recent school effects research, the dependent variable has been the gain in achievement, say between 10th and 12th grade, and this may cause problems. The commonly used model is

$$y(t_2) - y(t_1) = c_0^* + b^* y(t_1) + c_1^* x_1 + c_2^* x_2 \ldots . c_n^* x_n \qquad (6)$$

If the left-hand side is replaced by $y(t_2)$ alone, as in Equation (5), the estimates of the c_i coefficients are the same, but the coefficient for $y(t_1)$ becomes $b^* + 1$. Equation (6) is referred to as the regressor variable model in the psychometric literature, and the variant with only $y(t_2)$ as the dependent variable is known as the lagged model in the econometrics literature. These formulations, with $y(t_1)$ on the right-hand side, are usually seen as preferable to other models because they avoid the problems caused by the well-known unreliability of the change score, $y(t_2) - y(t_1)$, and they capture the regression toward the mean that otherwise may seriously bias other coefficients correlated with $y(t_1)$. Interpretations of the model vary. Chubb and Moe (1990) saw $y(t_1)$ as a measure of ability affecting level of achievement at Time 2. Coleman and Hoffer (1987) took the model as one for gains in achievement. Equation (6) is identical to Equation (5) and therefore also represents the estimation equation for the model of learning we have described previously with the model for change in $v(t)$ over time given in Equation (4).

The regressor variable method (or the lagged model) is not free of problems. If it is used to estimate the difference in achievement gain for two groups with different initial levels of achievement, such as Catholic and public schools, measurement error in y will produce under adjustment for the initial difference between the two groups (see Judd & Kenny 1981, Chapter 6; and Willett 1988 for further explanation). Allison (1990) proposed the estimation of models of the change score without $y(t_1)$ as a right-hand-side variable in order to avoid the bias caused by measurement error in $y(t_1)$. When the effect of the x_i variables occurs after $y(t_1)$ is obtained, Allison showed that the measurement error problem disappears.

Allison's (1990) argument amounts to estimating the fixed-effect (or first difference) model from the econometrics literature. This model is obtained by assuming a model for the levels of achievement at Time 1 and Time 2 and then subtracting the two models. This results in the equation

$$y(t_2) - y(t_1) = b_0 + b_1\Delta x_1 + b_2\Delta x_2 \ldots b_n\Delta x_n \qquad (7)$$

where b_i are the coefficients to the x variables that change between Time 1 and Time 2. Constant variables, including unmeasured variables that remain constant over the period, disappear and are eliminated as sources of bias. Such fixed-effects models have become very popular for the treatment of panel data. Following Allison, they may also seem appropriate for the estimation of models for achievement gains, removing the problem caused by measurement error in $y(t_1)$ or by unmeasured variables correlated with $y(t_1)$.

It is important to note that the fixed-effect or the change score model assumes that changes in y_t are produced by changes in independent variables. In other words, the fixed-effect model assumes that academic achievement only increases with an increase in some determinant of learning. This is not a reasonable assumption for the type of variables and processes usually investigated in school effects research. Measured family background variables usually do not change markedly over time. In any event, to the extent that the background variables are measures of stable ability differences, the mechanism by which they affect learning is one where the level of ability, or cultural resources in the home, affects gains in achievement. This is also what the learning model, suggested previously, assumes. Furthermore, the opportunities for learning provided by a school may be considered stable characteristics of schools.

The statistical advantage of the change score or fixed-effect model holds when ordinary least squares are used. The lagged models, of which Equation (6) is an example, have generated a rich literature in econometrics (e.g., Greene, 1993) in the context of time series models. Here the proposed solution to the biases caused by $y(t_1)$ is to use instrumental variables. The

lagged models do seem more appropriate for studying learning outcomes of schools than the fixed effect model. However, an instrumental variables approach to overcome the statistical problems with measurement error in the lagged variable has not yet made an appearance in the school effects literature.

CONCLUSION

No body of sociological research has generated as much controversy as the school effects research conducted by James S. Coleman. The policy implications have been unpopular with many of the interest groups associated with schools. Because the conclusions were based on some of the largest quantitative research enterprises conducted, much effort has been invested in finding methodological faults and problems with the research. We have reviewed some of these methodological problems, and clearly some conclusions are dubious and are based on procedures and methods that now seem inadequate. Nevertheless, we find that the most serious problems are conceptual. School effects research has paid very little attention to the issue of how school effects are brought about, and when the research has proposed such mechanisms, as in the research on Catholic and on public schools, they have not been empirically established.

We propose the use of a nonlinear specification of how the learning process depends on student endowments and the main activity of schools, teaching. The model is based on a simple theory of how learning outcomes are produced, a theory that has important implications for how one should go about conducting school effects research. However, we are doubtful that such an approach will gain wide acceptance in future school effects research.

There is a serious dilemma for policy oriented sociological research here. Sociologists have few well-specified theories that result in acceptable mathematical models for processes, and those theories that do exist, such as the one we outlined above, are unlikely to be agreed upon. This means that sociologists are unlikely to propose that a policy recommendation be based on models that deviate from the standard linear statistical models we are taught to use by statisticians. Statisticians do not have sociological theories. They, therefore, propose models that are parsimonious. This is a virtue, but the result may be an incorrect representation of the processes under study. The simple additive model for academic achievement from our perspective is a poor theory. However, it is the only theory likely to be believed by the research community at large and the policymakers informed by this community. Sociologists therefore must choose between believing their own theories while accepting the difficulty of getting the models they imply accepted, or instead ignoring the implications of their own theories in order to provide standard statistical descriptions.

ACKNOWLEDGMENT: The research was supported by a grant from the American Educational Research Association which receives funds for its AERA Grants Program from the National Science Foundation and the National Center for Education Statistics (U.S. Department of Education) under NSF Grant #RED-9452861. Opinions reflect those of the authors and do not necessarily reflect those of the granting agencies. We are indebted to Susan Dumais, Deborah Hordon, Gabriella Gonzalez, and Shannon Hodge for research assistance.

REFERENCES

Alexander, K. L., & Pallas, A. M. (1983). Private schools and public policy: New evidence on cognitive achievement in public and private schools. *Sociology of Education, 56,* 170–182.

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology, 20,* 93–114.

Angrist, J. B., & Krueger, A. B. (1991, November). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics, CVI*(4), 980–1014.

Barr, R., & Dreeben, R. (1983). *How schools work.* Chicago, IL: Chicago University Press.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children,* 5, 25–50.

Berliner, D. C., & Biddle, B. J. (1997). *The manufactured crisis.* White Plains, NY: Longman.

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

Cain, G. G., & Watts, H. W. (1970). Problems in making policy inferences from the Coleman Report. *American Sociological Review, 35,* 228–242.

Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools.* Washington, DC: Brookings.

Coleman, J. S. (1961). *The adolescent society.* New York, NY: The Free Press.

Coleman, J. S. (1972). The evaluation of "equality of educational opportunity." In F. Mosteller & D. P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 146–167). New York, NY: Random House.

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

Coleman, J. S., Kelly, S., & Moore, J. (1975). *Trends in School Segregation, 1968–1973.* Washington, DC: The Urban Institute.

Coleman, J. S., & Hoffer, T. (1987). *Public and private schools: The impact of communities.* New York: Basic Books.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement.* New York: Basic Books.

Cremin, L. A. (1951). *The American common school.* New York: Knopf.

Goldberger, A. S., & Cain, G. G. (1982, April/July). The causal analysis of cognitive outcomes in the Coleman, Hoffer and Kilgore Report. *Sociology of Education, 55,* 103–122.

Goodlad, J. I. (1983). *A place called school: Prospects for the future.* New York: McGraw-Hill.

Greene, W. (1993). *Econometric analysis. 2nd Edition.* Englewod Cliffs, NJ: Prentice-Hall.

Hanushek, E. A. (1971, May). Teacher characteristics and gains in academic achievement. *American Economic Review, 60,* 280–288.

Hanushek, E. A. (1986, September). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24,* 1141–1177.

Hanushek, E. A. (1994a). *Making schools work.* Washington, DC: The Brookings Institution.

Hanushek, E. A. (1994b). Money may matter somewhere: A response to Hedges, Laine, and Greenwald. *Educational Researcher, 23,* 5—8.

Hanushek, E. A. (1996). A more complete picture of school resource policies. *Review of Educational Research, 66,* 397–409.

Hanushek, E. A., & Kain, J. F. (1972). On thevValue of "Equality of Educational Opportunity" as a guide to public policy. In F. Mosteller & D. P. Moynihan (Eds.), *On Equality of Educational Opportunity* (pp. 69–115). New York: Random House.

Hauser, R. M. (1978). On "A reconceptualization of school effects." *Sociology of Education, 51*(1), 86–73.

Heckman, J. J. (1979). Sample bias as a specification error. *Econometrica, 47,* 153–162.

Heckman, J. J., & Neal, D. (1996). Coleman's contribution to education: Theory, research styles and empirical research. In J. Clark (Ed.), *James S. Coleman* (pp. 81–102). London: Falmer Press.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994, April). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher, 23,* 5–14.

Heyns, B., & Hilton, T. L. (1982, April/July). The cognitive tests for high school and beyond: An assessment. *Sociology of Education,* 55, 89–102.

Hoaglin, D. C., Light, R. J., McPeek, B., Mosteller, F., & Stoto, M. A. (1982). *Data for decisions.* Cambridge, MA: Abt Books.

Jencks, C. S., & Phillips, M. (1998). *The Black-White Test Score Gap.* Washington, DC: Brookings Institution Press.

Judd, C. M., & Kenny, D. A. (1981). *Estimating the effects of social interventions.* New York: Cambridge University Press.

Kerckhoff, A. C. (1993). *Diverging pathways: Social structure and career decisions.* New York: Cambridge University Press.

Lightfoot, S. L. (1983). *The good high school: Portraits of character and culture.* New York: Basic Books.

Mosteller, F. (1995). The Tennessee study of classs size in the early school grades. *The Future of Children,* 5, 113–127.

Mosteller, F., & Moynihan, D. P., (Eds.). (1972). *On equality of educational opportunity.* New York: Random House.

Murnane, R., Newstead, S., & Olsen, R. J. (1985). Comparing public and private schools: The puzzling role of selectivity bias. *Journal of Business and Economic Statistics, 3,* 23–35.

Murnane, R. J., & Phillips, B. (1981, March). What do effective teachers of inner-city children have in common. *Social Science Research, 10,* 83–100.

National Center for Educational Statistics. (1993). *Education in the states and nations: Indicators comparing U.S. states.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Noell, J. (1982, April/July). Public and Catholic schools: A Reanalysis of "Public and Private Schools." *Sociology of Education, 55,* 123–132.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The shopping mall high school: Winners and losers in the educational marketplace.* Boston: Houghton Mifflin.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70,* 41–55.

Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development.* New York: Holt, Rinehart and Winston.

Schultz, T. (1961, March). Investment in human capital. *American Economic Review, 51,* 1–17.

Sizer, T. R. (1984). *Horace's compromise: The dilemma of the American high school.* Boston: Houghton Mifflin.

Smith, M .S. (1992). Equality of educational opportunity: The basic findings reconsidered. In F. Mosteller & D. P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 230–342). New York: Random House.

Sørensen, A. B. (1987). The organizational differentiation of students in schools as an opportunity structure. In M. T. Hallinan (Ed.), *The social organization of schools: New conceptualizations of the learning process* (pp. 103–130). New York: Plenum.

Sørensen, A. B. (1996). Educational opportunities and school effects. In J. Clark (Ed.), *James S. Coleman* (pp. 207–225). London: Falmer Press.

Sørensen, A. B., & Hallinan, M. T. (1977). A reconceptualization of school effects. *Sociology of Education, 50,* 522–535.

Sørensen, A. B., & Hallinan, M. T. (1986, Winter). Effects of ability grouping on growth in academic achievement. *American Educational Research Journal, 23*(4), 519–542.

Westbury, I. (1992). Comparing American and Japanese achievement: Is the United States really a low achiever? *Educational Researcher, 21*(5), 18–24.

Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of Research in Education, 1988–89,* 345–422.

Willms, J. D., & Jacobsen, S. (1990). Growth in mathematics skills during the intermediate years: Sex differences and school effects. *International Journal of Educational Research, 14,* 157–174.

Winship, C. & Mare, R. D. (1992). Models for Sample Selection Bias. *Annual Review of Sociology, 18,* 327–350.