DOCUMENT RESUME

ED 264 212 SP 026 833

AUTHOR Good, Thomas L.; Brophy, Jere E.
TITLE School Effects. Occasional Paper No. 77.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jul 85
CONTRACT 400-81-0014
NOTE 109p.; Table I may not reproduce clearly.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Academic Achievement; *Educational Environment; Educational Objectives; Elementary Secondary Education; Improvement Programs; *Research Needs; *School Effectiveness

ABSTRACT

This paper reviews the literature on school effects and school improvement. The literature reveals that some schools are much more effective than others at optimizing achievement and other student outcomes, and that schools judged successful by these criteria tend to have in common certain leadership, goal commitment, and school atmosphere factors that can form a basis for school improvement efforts, at least in urban schools. However, the available data base is very limited, suggesting the need for more extensive data on more aspects of schooling collected on a greater variety of schools, and the need for caution to avoid developing overly rigid or generalized prescriptions from the data that do exist. (Author)

Occasional Paper No. 77

SCHOOL EFFECTS

Thomas L. Good and Jere E. Brophy

Occasional Paper No. 77

SCHOOL EFFECTS

Thomas L. Good and Jere E. Brophy

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824-1034

July 1985

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Education. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-81-0014)

Institute for Research on Teaching

The **Institute for Research on Teaching** was founded at Michigan State University in 1976 by the National Institute of Education. Following a nationwide competition in 1981, the NIE awarded a second contract to the IRT, extending work through 1984. Funding is also received from other agencies and foundations for individual research projects.

The IRT conducts major research projects aimed at improving classroom teaching, including studies of classroom management strategies, student socialization, the diagnosis and remediation of reading difficulties, and teacher education. IRT researchers are also examining the teaching of specific school subjects such as reading, writing, general mathematics, and science, and are seeking to understand how factors outside the classroom affect teacher decision making.

Researchers from such diverse disciplines as educational psychology, anthropology, sociology, and philosophy cooperate in conducting IRT research. They join forces with public school teachers, who work at the IRT as half-time collaborators in research, helping to design and plan studies, collect data, analyze and interpret results, and disseminate findings.

The IRT publishes research reports, occasional papers, conference proceedings, a newsletter for practitioners, and lists and catalogs of IRT publications. For more information, to receive a list or catalog, and/or to be placed on the IRT mailing list to receive the newsletter, please write to the IRT Editor, Institute for Research on Teaching, 252 Erickson Hall, Michigan State University, East Lansing, Michigan 48824-1034.

Abstract

This paper reviews the literature on school effects and school improvement. The literature reveals that some schools are much more effective than others at optimizing achievement and other student outcomes, and that schools judged successful by these criteria tend to have in common certain leadership, goal commitment, and school atmosphere factors that can form a basis for school improvement efforts, at least in urban schools. However, the available data base is very limited, suggesting the need for more extensive data on more aspects of schooling collected on a greater variety of schools, and the need for caution to avoid developing overly rigid or generalized prescriptions from the data that do exist.

SCHOOL EFFECTS[1]

Thomas L. Good and Jere E. Brophy[2]

This paper concerns a topic which scholarly journals and the popular press have afforded considerable attention during the past five years. Although issues related to the effects of school on student achievement are salient today, interest in effective schools is recent. In 1970, Biddle noted that little systematic study of school process had occurred, and the topic of effective schools received scant attention in the *Second Handbook of Research on Teaching* (Travers, 1973). Despite considerable interest in the issue today and numerous recently-published articles, relatively little process data describes what takes place in schools generally or how schools that do influence student progress positively differ from those with less positive impact.

Following the publication of the well-known Coleman, Campbell, Hobson, McPartland, Mood, Weinfield, and York report in 1966, many researchers attempted to relate school inputs to school outputs; however, this research ignored what took place in schools. Extensive reviews of the input-output literature (see for example Averch, Carroll, Donaldson, Kiesling, & Pincus, 1974) suggest that these studies fail to provide any *consistent* evidence for a

---

[1]This paper is a chapter in M.C. Wittrock (Ed.), *Handbook of research on teaching*, third edition. New York: Macmillan, in press.

[2]Thomas L. Good is a research scientist with the Center for the Study of Social Behavior, University of Missouri-Columbia, and a professor of curriculum and instruction. Jere E. Brophy is co-director of the IRT.

The authors wish to acknowledge the typing support provided at the Center for Research in Social Behavior, University of Missouri-Columbia, and to specifically thank Patricia Shanks for typing the chapter, finding references, and capably assisting in finalizing the manuscript. Gail Hinkel provided valuable editorial and content suggestions. They also want to thank three other colleagues for their careful review/critique of this chapter--Steven Bossert, Michael Cohen, and Marshall Smith.

relationship between general school resources and student outcomes (such as achievement). Consequently, this paper focuses on recent research that examines school process and its relationship to student outcomes. Readers interested in a historical overview of earlier research can find that elsewhere (e.g., Averch et al., 1974; Miller, 1983).

Research on school effects has not examined process variables in as much detail as research on teaching (see Chapters 10 and 13 in Wittrock, in press). Still, the 1970s were fruitful years of progress in advancing knowledge of schools and their effects on pupils.

In this paper we describe here school effectiveness research in the 1970s and 1980s--what is known about public schools' effectiveness in promoting the *average academic achievement* of students they serve--because most research in the last decade has examined average effects of schools on students. Space limitations prevent the review of important related topics (e.g., private schools, desegregation, effects of schooling on mainstreamed students).

Although we have organized the paper to facilitate the task of reviewing extant literature, we do not feel that extant research has explored the effects of schools on students in any systematic fashion. Indeed, later in the paper we raise questions about the validity of extant research and suggest other directions for future research. At this point, suffice it to say that student achievement on standardized test scores cannot be equated with effectiveness per se. Schools are asked to influence many aspects of students' behavior and attitudes. Hence, information about school effects on narrow measures of student achievement is relevant and interesting but only one of many dimensions of schooling that contribute to assessing the general concept of effectiveness. We use the term "effective schools" frequently in this paper despite our earlier qualifying remarks about the limitations of

extant findings because the literature is replete with references to effective schools. However, we will use this term in a very restricted sense--to refer to schools that have obtained relatively high amounts of average student achievement.

Do Schools Make a Difference in Student Achievement?

Some argue on the basis of the Coleman et al. (1966) study and subsequent input-output research that studying schooling is pointless because the effects of schooling on student achievement are minor at best, and the topic frequently debated. For example, Jencks, Smith, Ackland, Bane, Cohen, Gintis, Heyns, and Michelson (1972) and Rowan, Bossert, and Dwyer (1983) agree that schooling accounts for but a small percentage of variation in student achievement, but disagree markedly on the importance of such effects. Some claim that statistical procedures seriously underestimate the effects of schooling on achievement. For example, Madaus, Kellaghan, Rakow, and King (1979) argue that school-related variance in student achievement is large when one focuses on subjects actually taught in schools. Still, most research indicates that family background variables affect student achievement more than school variables, although researchers and reviewers generally agree that (a) school effects are still important and (b) school influence on student achievement is generally underestimated.

Rutter (1983) argues that the impact of schooling on achievement is underestimated because of (a) the outcome variables measured, (b) the predictor variables measured, and (c) the extent of variation of predictor variables. He notes that the teaching of verbal skills (as measured in early input-output studies) is not the main objective of schooling. Accordingly, estimates of school effects have sometimes been based on measures that bear little relationship to what most schools attempt to teach. Several recent studies show that school variables account for more variance in pupil

achievement in specific curriculum-based subjects and on norm-referenced tests (as opposed to general tests of verbal facility), though school variables still constitute only a small amount of the total variance (Brimer, Madaus, Chapman, Kellaghan, & Wood, 1978; Madaus et al., 1979; Postlethwaite, 1975).

Predictor variables also affect results of research on school effects. Most surveys consider a narrow range of school variables and focus on financial or physical resources rather than the internal social life of schools. Growing evidence shows, however, that social and instructional variables, rather than financial variables, account for important variation among schools.

If research uses some average measure of an entire school, all children at that school will necessarily receive the same (school) score. As Rutter (1983) points out, this procedure involves the misleading assumption that all students in any school receive the same school experiences. Consequently, results based on such statistical analyses in many instances will underestimate the size of school effects.

However, Rutter notes that, other things being equal, a predictor variable with a wide range will account for a higher proportion of the variance than a predictor with a narrow range. This is because schools tend to be more homogeneous (with respect to certain characteristics like use of language and techniques used to discipline children) than families, and because the difference between the best and worst schools is likely to be far less than that between the best and worst homes. Aggregating data with the school as the unit of analysis means averaging across students from widely contrasting family types and averaging across teachers who vary in instructional effectiveness. It also means that variance across schools will be lower than the variance across both students and teachers.

Absolute Effects of Schooling

In terms of population variance accounted for, family variables will usually have a greater effect than school variables, but schools will not necessarily have less influence than families on achievement. Rutter offers the following hypothetical example. Assume that the outcome variable is pupils' achievement in Sanskrit, that books on Sanskrit are available only to the teachers at the school, and that all schools are equally good at teaching Sanskrit. Because Sanskrit can be learned only at school, schools are necessarily the only direct influence on Sanskrit achievement. But because of variation in pupils' ability to learn (as a result of both genetic and environmental influences), some children will achieve much higher levels in Sanskrit than other children. However, because all schools teach Sanskrit equally well, schooling would account for *none* of this individual variation. In short, in this situation schools would have no measured effect on Sanskrit achievement in spite of the fact that all Sanskrit was necessarily learned only as a result of schooling.

Existing claims about the effects of schools on achievement are drawn from societies where education is compulsory (Europe, North America, Japan); hence, all students are exposed to teachers, texts, and curriculum assignments. That all students have such advantages is a factor that leads to *underestimation* of the effects of schools on achievement. Heyneman and Losley (1983) examined the effect of primary-school quality on academic achievement in 29 high- and low-income countries. These researchers found that in low-income countries (where schooling is not compulsory) the predominant influence on student learning is the quality of the schools and teachers to which children are exposed (family background characteristics are considerably less important).

Heyneman and Losley argue that the skepticism of critics of U.S. schools about the efficacy of educational investments appears to be unwarranted or at least premature. They note that when international data are used in input-output regression models (like those used by Coleman et al., 1966 and others), *school and teacher quality* appear to be the major influences on student learning around the world.

Relative Effects of Schooling

Heyneman and Losley's data suggest that attending school per se has an absolute effect that tends to be masked in advanced societies because virtually all students attend school. However, even in a relative sense, some forms of schooling appear to have important effects. For example, Rutter, Maughan, Mortimore, Ouston, and Smith (1979) found that after adjusting for intake characteristics, children at the most successful secondary school passed an average of four times as many exams as children at the least successful school. Also, children in the bottom 25% on verbal ability in the most successful school on average passed as many exams as children in the top 25% on verbal ability at the least successful school.

Brookover, Beady, Flood, Schweitzer, and Wisenbaker (1979) found that children in "successful," white, elementary schools obtained academic achievement scores an average of about one standard deviation above those in the "unsuccessful," white, elementary schools matched for intake. The difference between the successful and unsuccessful black schools was even greater. However, as Rutter notes, these were average scores and within *all* schools there are children with both superior and inferior achievement. Moreover, these differences are among schools at the extremes of the achievement range. These data clearly show that the effects of improving the quality of the worst

schools are likely to be great enough to be of considerable practical importance.

Considering that schools can have important effects on students' achievement, we now examine research of the 1970s and 80s on school effects. What is known about how schools vary in instructional, organizational, and social processes and how does such variation relate to student performance?

Studies of Unusually Effective Schools

Much research in the late 1960s and early 1970s suggests that differences in school resources and practices do not relate to variations in student achievement as measured by standardized achievement tests. Klitgaard and Hall (1974) argue, however, that one of several methodological problems in earlier research (e.g., Coleman et al., 1966) was that previous studies of school effectiveness measured only general school effects (the average effect of all schools in a sample) on measurable student outcomes. They state that even if extant data are accepted (along with questionable assumptions and procedures), there could be some unusually effective individual schools. Furthermore, some unusually *ineffective schools* can also be masked when data are reported only in *group averages*, but Klitgaard and Hall did not explore this possibility.

Student progress clearly varies from school to school, but the most important issue is whether variation in achievement among schools is affected by school process or whether this variation can be explained completely in terms of student factors (e.g., aptitude). The question posed by Klitgaard and Hall is crucial and addresses similar issues examined in recent process-product studies of teacher effectiveness (see chapter by Brophy & Good in Wittrock, in press). If some meaningful variation exists in performance among schools, then there is reason to believe that student performance in many schools can be improved.

Do Effective Schools Exist?

Klitgaard and Hall. We have discussed general issues and now examine some of the literature. The first study, Klitgaard and Hall (1974), is important for historical as well as substantive reasons. It was the first rigorous, large-scale attempt to find effective schools.

To determine whether unusually effective schools exist, Klitgaard and Hall operationally defined effectiveness as student performance on standardized reading and mathematics achievement tests. They analyzed three data sets: the 1969-1970 and 1970-1971 Michigan assessment of fourth and seventh grades (drawn from 90% of the state's public schools); scores from Grades 2-6 from 1967-1971 in New York City; and test scores from the Project Talent high schcol data of 1960.

The investigators examined histograms of the residuals from a regression of achievement scores on background factors and studied a series of distributions of residuals. The investigators calculated the cumulative total of these residual achievement scores for each school across years and tested to see if some schools were one standard deviation above the mean more often than chance would predict. They report that of the 161 Michigan schools that reported scores for all eight grade-year-test combinations, 15 were one standard deviation (or more) above the mean six out of eight times (less than one time would be expected by chance). Thus about 9% of the schools in the ssmple increased student achievement from the 50th to the 72nd percentile (with student background factors controlled statistically).

Klitgaard and Hall report that many of the outstanding schools were rural. Nevertheless, when rural schools were excluded from the analyses, they still found schools in which students consistently achieved at higher-than-average levels.

These investigators also examined their data for unusually effective school districts. Among 627 districts studied in New York, 30 were one standard deviation above the mean at least five out of eight times (less than four districts would be expected by chance). Finally, they found little evidence of unusually effective grade levels.

Although the data support the contention that some unusually effective schools exist, the results basically support previous research that indicates that the effects of schools are small after nonschool factors (socioeconomic status (SES), aptitude) are controlled. The high-achieving schools identified represented only from 2% to 9% of the sample. These schools were clearly unusual and had relatively more achievement than schools with comparable populations; however, whether they were effective depends upon one's definition of effectiveness (What performance measure is appropriate and how high does a school have to be on that measure[s] to be considered effective?).

This informative study supplies plausible data showing that at least some schools are more effective than would be predicted by chance. However, one issue left unexamined in this study is to what extent a high-achieving school is equally effective for *all* students enrolled. That is, a high-achieving school may be so classified because it is unusually effective with one group of students, even though it has no special effects on the remainder of the school population. High-achieving schools could even have detrimental effects for certain students or small groups of students. This issue illustrates how relative and elusive the concept of effectiveness is. We explore this important definitional issue later in this paper.

Initial school effects studies (input-output studies) thus examined the average effects of schools, and more recent studies concern the effects of individual schools (or groups of schools defined as effective or ineffective)

on average student performance. We will argue later that researchers should now consider the effects of schools on specific types of students.

Also, it is probably easier to identify *in*effective schools, and data from only the top half of the distribution (of residual gain scores being examined) probably underestimate the stability of school effectiveness. To our knowledge researchers have not examined ineffective schools in this way, and this seems to us an important consideration.

A Study of School Processes

Weber. In addition to being among the first to search for effective schools, Weber (1971) conducted one of the earliest studies designed to identify the *processes* operating in effective inner-city schools. Weber believed that some inner-city schools had more positive effects on student achievement than others. He rejected the assumption that low student intelligence and a lack of funds sufficiently explained low-income children's failure to learn. He tested these proposals in the third grade.

To identify potentially effective inner-city schools Weber asked reading specialists, publishers, and school officials for nominations. He kept the nomination process open for over a year. He did not intend to find all of the inner-city schools that were successful in beginning reading instruction, but rather to identify enough schools so that he could describe and analyze several representative, successful schools. A total of 95 schools were nominated. Of these, only 69 seemed to be serving nonselective public school populations. To each of these 69 schools he wrote a letter asking if the principal believed that the school met the criteria (inner-city school; successful in teaching reading) and if the school desired an independent evaluation of reading achievement. Some principals did not respond; others replied that they were not inner-city schools or that, by the criteria employed, they

were not successful in beginning reading instruction. A number of principals refused to participate when the nature of the independent evaluation was detailed. Weber ultimately visited 17 schools in seven large cities.

Weber conducted an independent evaluation of reading achievement to eliminate the possibility (however remote) that any student coaching and manipulation of scores might occur. Because he was interested in testing the ability of poor children to read words that they already understood, Weber devised a test made up entirely of words that he thought students would understand. The test was different from nationally standardized reading achievement tests in that it did not evaluate breadth of aural vocabulary or ability to take multiple-choice tests, but rather the ability to read simple American English. The final test contained 32 items and could be administered in 15 minutes.

The 17 schools participating in the project were visited for two or three consecutive days between January and June of 1971. Six of the 17 schools that were observed and tested met the inner-city criteria but not the reading-success criteria. Seven schools met the reading-success criteria but not the inner-city criteria. Four met both criteria, and Weber argues that these nonselective public schools in the central areas of large cities were attended by very poor children. The third-grade median reading achievement scores of these four schools equalled or exceeded the national norm, and the percentages of nonreaders were unusually low for such schools.

What, then, were the factors that distinguished these four inner-city schools from other schools that did not teach beginning reading as well? To his credit, Weber notes that it is impossible to be certain of the answer because schools do many things differently and it is difficult to determine which practices are responsible for high pupil achievement. However, he cites

eight characteristics of these schools: strong leadership, high expectations, good atmosphere, emphasis on reading, additional reading personnel, use of phonics, individualization, and careful evaluation of pupil progress. Weber does not mean individualization in the narrow sense of having each child work at a different level; rather, the term implies concern for each child's progress and a willingness to modify a child's assignments if necessary.

Some of the factors *not* associated with achievement were small class size, achievement grouping (in one school all classes were heterogeneous), and physical facilities (not one of the four school buildings was modern; two were noticeably old). Two years later Weber revisited two of the four successful schools and found that one of them continued to be effective (and had even improved somewhat); however, the other school had deteriorated notably and was no longer effective. Hence, the conditions of effective schools may be only temporary, and as principals, teachers, and student cohorts change so too may the level of school effectiveness.

It is unfortunate that Weber does not explain why this one school deteriorated. Considering that many schools are presently working to increase their effectiveness, it would be important to identify factors that were associated with the decline (e.g., change in student and/or teacher population) in student achievement in this school.

The observational data in this study provide only limited information; they yield hypotheses for future testing, not clear guidelines for effective schools. To begin with, other investigators may not have described these schools as Weber did (the same behaviors, policies, and standards may constitute various types of leadership or expectations to different people). The instrument Weber used is brief, and even though it is appropriate in other respects, a 15-minute test provides limited data for assessing student

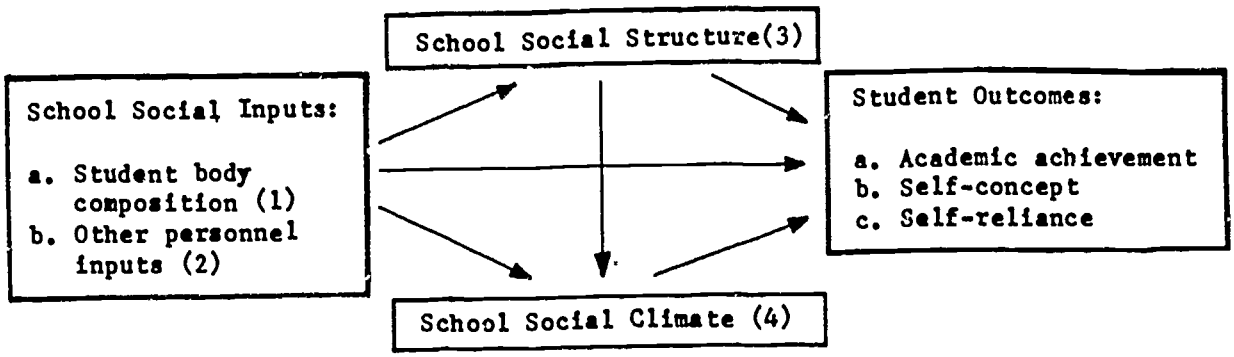
knowledge. *Since the study did not include observations of ineffective schools or average schools, it is very difficult to assess whether the factors identified by Weber have any true relationship with school achievement.* Furthermore, Weber's report does not comprehensively indicate which aspects of schooling did not differentiate effective schools. Still, considering the exploratory nature of his investigation, the lack of systematic reporting of what was observed (and the procedures for such observations) is understandable. Nevertheless, this places limitations on the confidence with which one can rely on the process measures used in the study. Weber was very successful, however, in stimulating others to explore the issue of how schools make a difference in student achievement.

Two Recent Empirical Studies of School Effects

In this section of the paper we discuss two of the most rigorous and salient process-product studies of school effectiveness. Although there have been several comprehensive studies, we limit discussion here to only two of these more important and more salient studies so that we can assess more adequately the particular research strategies used and see more fully the strengths and weaknesses of extant research. These studies are somewhat similar to the research paradigm Klitgaard and Hall (1974) used; however, after identifying outlier schools, the researchers attempted to explain how more and less effective schools varied in school routines and classroom practices.

A Recent Study of Effective Schools: Brookover et al. (1979)

Brookover et al. (1979) argue that the social system of a school influences the role definitions, norms, expectations, values, and beliefs that students internalize and that such socialization affects students' achievement, academic self-concepts, and other affective responses (see Figure 1).



(1) measured by mean school SES and percentage white.
(2) measured by standard scores of shcool size, average daily attendance, professionals per 1,000 students, average years teaching experience, average tenure in school, percentage of teachers with advanced degrees, and mean teacher salary.
(3) measured by teacher satisfaction with school structure, parent involvement, differentiation in student programs, principal's report of time devoted to instruction, open-closed classroom.
(4) measured by 14 variables derived from student, teacher and principal reports of the norms, expectations, and feelings about the school.

From Brookover et al., 1979.

Figure 1. General model of school social system variables with hypothesized relation to student outcomes.

The model suggests that the behavior students learn and their achievement will vary among schools and that this variation can be explained by differences among schools in inputs (quality of teachers and students), social structure, and climate. As Figure 1 demonstrates, Brookover et al. believe that the initial characteristics of teachers and students affect student outcomes. However, the quality of teachers and students is modified by school structures, processes, and beliefs.

Brookover et al. studied 68 schools drawn from a state pool that represented a random sample of Michigan fourth- and fifth-grade students. Sixty-one of these schools had populations that were more than 50% white. (For some analyses the white school sample was divided at the median of the SES distribution so that high- and low-SES white schools could be compared.) The black

school sample was composed of seven schools from the state sample whose populations were more than 50% black along with 23 other majority black schools randomly selected from the population of schools with majority black populations. Data were obtained from (a) the Michigan School Assessment Reports, (b) questionnaires administered to fourth- and fifth-grade students, (c) questionnaires administered to teachers, and (d) school principals. Complete questionnaires can be found in Brookover et al. (1979).

Input variables. The major school input variables assessed were (a) social composition of the student body, (b) school social structure, and (c) school climate. Important *outcome* variables studied were (a) student achievement, (b) self-concept of academic ability, and (c) self-reliance. *Social composition* in this study was defined as the mean socio-economic status of the school and the percentage of white students in the school. Other school input measures included the size of the student population, the average daily attendance of students, the number of staff per 1,000 students, and several teacher characteristics (experience, etc.).

The five factors defining the *social structure* of the school were (a) parent involvement, (b) differentiation of student programs (e.g., degree of ability grouping, use of students' interests in planning instruction), (c) openness of classroom organization (how often students talk and work together; how frequently seat assignments are changed), (d) time allocation (time allotted to academic, social, and administrative tasks), and (e) staff satisfaction with school structure.

In this study *school climate* was defined as the composite of norms, expectations, and beliefs about the school social system as participants perceived it.

Output variables. The measure used for *academic achievement* was the average percentage of students who mastered each of the 40 objectives in the Michigan School Assessment Test administered in the fall of 1974. Nineteen achievement objectives had been established for reading and 30 for mathematics. Brookover et al. used the mean percentage of all reading and mathematics objectives mastered to reflect achievement on a total of 245 questions.

The *self-concept of academic ability* scale focused on one aspect of self-concept--perception of self as student. The other self-concept measure was student perception of self-reliance (the extent to which students could and wanted to complete tasks or to solve problems on their own). Many of the social structure, climate, and outcome variables were assessed by questionnaires (for complete details see Brookover et al., 1979).

Interrelations among input variables. Three of the five variables used to define social structure were positively and significantly correlated with social composition and other input variables and intercorrelated with each other (the variable "degree of differentiated programs" did not correlate in any meaningful way with other input or structural variables). Brookover et al. note that although school structure variables were related to each other and to composition measures, they were not entirely dependent on composition nor were they merely different measures of the same variable.

Some measures of school climate were highly correlated with student body composition. The authors note that such high correlations make it difficult, if not impossible, to separate the effects of climate from those of school composition on achievement. For example, students' sense of academic futility (a set of items generally identified as measuring sense of control) correlated .87 with the composition of a school's population. With the exception of

differentiated programs, all input, structure, and climate variables were interrelated, at least to some degree.

Interrelations among dependent variables. Perhaps most important, a negative correlation of -.55 was obtained between *mean* school achievement and *mean* self-concept of academic ability. Thus students in lower-achieving schools actually averaged higher self-concept scores than students in higher-achieving schools. In previous research, however, Brookover et al. (1962, 1965, 1967) had found that *individual* student self-concept of ability scores correlated positively with student achievement. Here, in the 30 majority black schools essentially no correlation existed between self-concept and achievement (.004), and the same general pattern was obtained in the majority white schools (.04). There was a negative relationship between self-concept and achievement (-.23) in the white, low-SES schools. In general, high achievement did not correlate with high student self-concept or self-reliance. Thus it seems that at least in some schools different outcomes are associated with various combinations of school climate and structure variables. However, other research reviewed by Rutter (1983) suggests that effective schools can positively affect multiple criteria simultaneously.

Relationships between school system variables and student achievement. The researchers note that except in the majority white schools (particularly in the high-SES sample), less than half of the between-school variance in achievement was uniquely attributable to either input or structure variables independently. In general, more of this variance was explained by complex school social-system characteristics. The researchers suggest that their analysis does not indicate which variable or set of variables in the social

system has the largest effect on achievement in all school situations, but they do argue that the small proportion of variance uniquely attributable to input variables strongly suggests that school variables are important factors affecting student achievement.

Case studies. Brookover et al. (1979) supplemented their statistical analyses with classroom observations and interviews with participants in four low-SES schools. Time spent in the schools ranged from three weeks to three months. Unfortunately, the researchers report no information about what was observed, how the data were collected, or the questions asked of the participants. Furthermore, there is no explanation concerning the number of teachers visited in each school and no serious attention to within-school variation. Hence, it is very difficult to assess these data and their implications for practice.

The reported criteria for selecting the four schools were (a) similar racial composition, (b) similar SES levels that were significantly lower than the mean SES level for the sample, (c) achievement scores above the sample mean in one school and below the sample mean in the other within each pair, and (d) urban location. The variables used to describe differences in high- and low-achieving schools were (a) time spent in instruction, (b) write-off (percent of students not expected to master curriculum), (c) teacher expectations, (d) reinforcement practices, (e) grouping procedures, (f) teaching games, (g) principal's role, and (h) commitment of teaching and administrative staff. These variables seem to represent important aspects of schooling, but the authors do not explain why they examined these variables instead of other potential ways in which schools might differ.

Although the findings reveal some general differences that distinguished high- from low-achieving schools, considerable variation existed in how

principals and teachers in high-achieving schools obtained their effects. In brief, teachers in the higher-achieving black and white schools spent more time on instruction. Schools allocated relatively few study periods, although students had small blocks of time for doing classroom assignments. Teachers spent as much time as was necessary to convey new concepts, problems, and so forth for each lesson. However, they did not seem to spend inordinate amounts of time on lessons students had already mastered. Observations showed that academic interaction between teachers and students occurred more frequently in the white high-achieving school than in the black high-achieving school. Overall, however, academic interactions between teachers and students were more frequent in the high-achieving white and black schools than in the lower-achieving white and black schools. In most of the classes observed in the lower-achieving black and white schools, students had a great deal of time to study, read, or play while the teacher attended to administrative duties such as grading papers. In sum, more time was allocated for instruction in the high-achieving schools, although more time was spent in active teaching in the higher-achieving white school than in its black counterpart. The differences in school processes across the four schools are summarized in Table 1.

Summary. In combination, the data from this study suggest that when teachers, principals, and students believe that academic achievement is possible, school climate is conducive to learning and student achievement is higher. However, these data are interrelated and correlational; hence, it is not possible to determine whether high achievement *preceded or followed* positive expectations, or if students' high expectations for learning *preceded or followed* those of adults in the school. The data therefore do not clearly indicate how schools should initially invest limited resources available for school improvement projects. The findings do suggest, however, that improvement is possible, and they indicate several variables that can be manipulated.

Table 1

Highlights of Four Case Studies: What Makes a Difference in Predominantely Black and Predominantely White High- and Low-Achieving Schools in Low-SES School Districts

	High-Achieving White School	Low-Achieving White School	High-Achieving Black School	Low-Achieving Black School
Time	Most of class time spent on instruction--except for one teacher 80-90% of time used.	Time spent on instruction varied between classrooms. In several classrooms only 10% of time spent on instruction. Several teachers had managerial problems. Many teachers who did not use low-level work to keep students involved.	Teachers did much teaching. While students were working the teachers were available for clarification and reteaching as necessary	Most teachers attempted to keep students busy but not a lot of productive task-relevant work was achieved. Very little academic interaction with students.
Write-Offs	Few students seen as destined to fail fail, as hopeless cases; no remedial programs.	Usually 2-3 students per class remained outside of learning process. However, most of "slower" students less involved in work and interactions with teachers.	Teachers felt vast majority of students were capable of mastering assigned materials. Only a very few students were seen as unlikely to make it. When one strategy didn't work, teachers were willing to try other strategies.	Teachers appeared to write off large numbers of children; large numbers of students were required to attend remedial classes and such classes were seen as "dumping grounds."
Teacher Expectations	Teachers expected students to work at grade level.	Expectations for student achievement were low in general but especially for students in slow reading tracks. Grade-level achievement not seen as a realistic goal for many students.	Teachers generally reported that they expected at least 75% of their students to master assigned work and that 75% would complete high school.	Teachers generally held low performance expectations for students and teachers were unwilling to assume responsibility for student learning.
Reinforcement Practices	Appropriate reward.	Teachers varied. Some teachers used appropriate reinforcement practices but several teachers were observed to use confusing and/or totally inappropriate reinforcement practices most of the time.	Teachers tended to use reinforcement patterns that were likely to encourage higher achievement.	Many teachers in the regular classrooms used reinforcement inappropriately, often telling students they had done well when, in fact, they had not.

	High-Achieving White School	Low-Achieving White School	High-Achieving Black School	Low-Achieving Black School
Grouping Procedures	No homogeneous grouping after third grade. Grouping to classrooms basically random.	Students grouped 1-6 for reading instruction. Only two groups per class--high and low. Mobility between groups very limited.	Students grouped on the basis of pretest on math and reading tests. Teachers appeared to teach with purpose of advancing students to higher groups when possible. Teachers suggested that the goal for a full year's achievement gain for all students was an academic floor and not a ceiling.	Extensive use of grouping; however, did not encourage advancing students to join higher groups and appeared to be more of a management than an instructional tool. Also, there was extensive assignment of "slow" students to remedial classes. The extensive grouping and regrouping appears to be disruptive.
Teaching Games	Reflected high expectations of teacher and appropriate reinforcement--emphasized team rather than individual learning.	No mention.	When games were used they tended to be team games. Games appeared to be used to reinforce an attitude of "try a little harder."	Teaching games seldom used in regular classes--used much in remedial classes, but tended to be individual games and not used in a way likely to stimulate achievement gain.
Principal's Role	Heavily involved in instructional issues. Provided instructional leadership. Assumed measure of responsibility for the educational functions of the school. Visited classrooms frequently.	Time shared between two buildings. Seemed to be a part-time administrator-part-time disciplinarian. Seldom visited classrooms; did not function as an educational leader. Often expressed low performance expectations for students.	Emphasized administrative duties although he did observe and critique teachers periodically. Note: the previous principal had been a more active educational leader.	Principal at this school was mainly as administrator and disciplinarian. Although the principal talked about the importance of student achievement, there was little pressure brought to bear by the principal on teachers to improve classroom performance.
Commitment	Commitment to high achievement--willing to make public announcements to one another and to parents that students could learn.	No explicit discussion.	Considerable interest in providing students with high-quality education. Much warmth directed toward individual students.	Teachers' behavior suggested that there was little they could do to increase student achievement.

BEST COPY AVAILABLE

In summary, the Brookover et al. (1979) study is a comprehensive and successful attempt to illustrate that school inputs do not predict student outcomes (achievement, self-concept, self-reliance) independently of school process. The case-study data suggest that schools with varying input resources will have differential effects on student achievement because of climate and structural features present in the schools. Furthermore, the authors find that climate variables (although highly correlated with input variables) explain as much variation in achievement as do input variables. Although the process data collected in this study do not yield definitive statements about school process in more and less effective schools, they do suggest that schools with comparable resources can have very different climates, and they provide a base on which future observational studies can build.

Another Modern Study of Effective Schools: Rutter et al. (1979)

In a three-year study of 12 secondary schools, Rutter et al. (1979) found that some urban, secondary schools were better than others in promoting students' academic and social success.

A survey of London 10-year-olds was initiated in 1970 when children were nearing the end of primary school. Assessments of intellectual level and reading attainment (as well as pupil behavior and family circumstances) were obtained for children attending primary school in one inner-London borough. All pupils were retested in 1974 at approximately age 14 during their third year of secondary school. Testing was concentrated in the 20 secondary schools which had taken the majority of the children tested in 1970 at age 10.

There were large differences between schools in terms of delinquency rates and achievement problems. However, to what extent were these secondary school differences merely a reflection of input differences (e.g., the

BEST COPY AVAILABLE

proportion of difficult children admitted from primary schools)? The data showed that there was not only *variation in output*, but there were also *substantial differences in input* that these 20 secondary schools had to work with (e.g., for boys, some schools admitted as few as 7% with behavior or reading problems; whereas others took as many as 48% to 50% with such problems).

Still, despite these large differences in input characteristics, after equating the pupil input of schools statistically, substantial and statistically significant differences existed between schools. Schools with the most advantaged students were not necessarily those with the best outcomes and schools that had students from similar backgrounds often had vastly different outcomes. Simply put, these data illustrate that pupil behavior and delinquency rates at age 14 could not be explained by family background variables or by pupils' test or questionnaire scores at age 10.

To study schools carefully with available resources, the investigators reduced the number of schools studied from 20 to 12. Three main measures were collected on each student: (a) verbal reasoning scores at the age of 10, (b) parental occupation, and (c) students' scores on a behavioral questionnaire completed by their primary school teachers.

The researchers examined the extent of change in intake in student population at each of the target schools over several consecutive years. In general, they found that the population at particular schools was relatively stable, so that those schools with the most advantaged *intakes* in 1971 maintained their advantaged status through 1974.

Process measures. The study neither tested a particular theory of schooling nor built on any preconceived ideas about which school processes are important. In general, the study examined processes derived from seven broad conceptual areas: academic emphasis; teacher actions in lessons, rewards and

punishments; conditions of learning for pupils; pupils' responsibilities and participation in the school; stability of teaching; and friendship-group organization. Data on processes were derived from interviews with teachers, pupils' responses to a questionnaire, and classroom observation.

Observational procedures. The main series of observations consisted of one week's observation in each school, in middle-ability, third-year classes. In all, 402 lessons were tape-recorded and coded; most of these lessons (312) were in academic subjects. The investigators coded the activities of both teachers and pupils in each lesson. The observation periods were grouped into coding sections that focused first on the teacher, then on selected individual children, and finally on the whole class. Each section lasted five minutes, and this pattern was repeated throughout the lesson. First, the coders recorded whether the teacher focused on the subject matter, pupils' behavior, or some other academic, social, or administrative activity. In addition, observers noted whether the teacher interacted with students and if so, whether with the whole class or with individuals. They also coded examples of praise or punishment and any marked expressions of warmth or negative feelings toward children.

Observers noted if students appeared to be engaged in tasks set by the teacher and the frequencies of other presumably less acceptable behaviors. The coders also selected five pupils at random and recorded whether they were working at the appropriate task. The researchers reported not only the occurrence of a behavior, but also the proportion of the class involved in it. For example, the average pupil on-task behavior was 81.5% of each lesson and about 75% of the teacher's time was spent on the subject matter of the lesson. In addition to structured, time-sampled observations, the investigators also

observed playground behavior, recording pupils' activities, and noting any physical violence between children and any unofficial sanctions by staff.

Outcome measures. The investigators used five outcome measures to assess the differential effects of individual schools: attendance, student behavior in school, examination success, employment, and delinquency. Here we discuss only the results associated with achievement, attendance, and student behavior.

The investigators report large differences in attendance rates across schools, even after variations in school inputs were controlled. Pupil behavior (e.g., late arrival, off-task behavior, disruptive behavior) also varied considerably across schools after differences in intake were controlled. The correlation between intake and behavior was only .27, suggesting that schools with more disruptive students (as defined in terms of students' behavior in the primary school) did not necessarily have the worst classroom behavior.

Although it was difficult to compare achievement because of variation in tests and curriculum emphases across schools and the complexity of the examination system, the researchers compared academic progress across the 12 schools. Even after controlling for variations in intake, a marked variation existed in achievement between schools. In general, school achievement was reasonably stable over two consecutive years (with the exception of one school). However, marked, positive variation in achievement was a characteristic of only two schools in the sample (see Purkey & Smith, 1983a).

Interrelations of outcomes. Table 2 illustrates the schools' relative rankings in terms of attendance, academic achievement, and desirable classroom pupil behavior. It shows that schools ranked high on any one of these measures tended to rank high on the other two, and schools ranked low on one

outcome also ranked low on the other two. However, Table 2 shows two glaring exceptions to this pattern. In one instance, students who behaved well

Table 2

A Comparison of School Effectiveness
Across Attendance, Academic Achievement,
and Behavior Outcome Measures

Attendance	Academic Achievement	Behavior
1	1	1
2	2	4
3	6	3
4	5	5
5	8	10
6	4	6
7	10	11
8	9	9
9	3	8
10	7	7
11	12	12
12	11	2

Note. 1 = highest, 12 = lowest

had poor *achievement* and attended school less often. In another school students made relatively high academic progress despite relatively low attendance levels and a high amount of inappropriate classroom behavior.

As Rutter and colleagues note, variations in outcome do not necessarily prove that outcomes are influenced by what happens in schools. However, these data do strongly suggest that something more than intake measures produced the differential outcomes, and one can plausibly infer that something in the school was at work. This argument would be stronger if the researchers could demonstrate that variations in school structure, school processes, and classroom behavior were consistently associated with variations in student

outcomes. We now discuss factors that consistently were found to be associated with more effective outcomes in the Rutter et al. study.

Process findings. An academic emphasis was associated with school achievement gains. For example, schools in which teachers assigned homework frequently (and teachers were monitored to be certain that they did so) tended to have higher achievement than schools that seldom assigned homework. Note, however, that the average time spent on homework (as reported by the pupils themselves) was *not* great in any of the schools and averaged only 15 to 35 minutes or so across schools. (Similar findings are also reported in classroom studies, e.g., Good, Grouws, & Ebmeier, 1983.)

These findings, of course, do not demonstrate how and why homework is associated with increased achievement. For example, it may well be that in addition to its practical value as distributed practice, homework also symbolically emphasizes a school's concern for academic progress and its expectation that pupils work independently.

Another aspect of academic emphasis concerns *teachers' expectations* for their pupils. Teacher expectations correlated positively and significantly with both attendance and academic outcomes. It could be argued that teachers are simply good judges of children's abilities. However, when teacher expectations were compared with student ability at intake, two schools in the bottom third with respect to academic expectations were in the top third with respect to initial student ability.

Furthermore, the proportion of the school week devoted to teaching was associated with greater student achievement. The time actually spent teaching varied among the 12 schools from 21.9 to 24.2 hours per week, and was positively correlated with pupil attendance. Taken together, the findings on academic emphasis suggest that students tend to make better progress both

behaviorally and academically in schools that focus on academic matters. This emphasis is reflected in a well-planned curriculum and in teachers' high academic expectations for students.

Teacher orientation. The amount of school time spent on the lesson topic varied from 65% to over 85%; nevertheless, time spent on the lesson topic was *not* significantly associated with academic success. An attentive, well-behaved class provides the opportunity for effective teaching and productive learning. What *use* teachers make of this opportunity, however, is crucial in determining what and how much children learn.

Teachers in more successful schools spent a higher proportion of their time interacting with the *class as a whole* than with individual pupils. Lessons in the successful schools more frequently included periods of quiet work when teachers expected pupils to work by themselves.

Frequent disciplinary interventions in a school were associated with increased pupil off-task behavior. Teacher behavior that results in many interruptions of lessons and involves constant checking and reprimanding may perpetuate student behavior problems. Conversely, schools where most lessons started promptly tended to have better outcomes and better student behavior. In general, the findings on teacher management behavior in this study are remarkably similar to those reported by Kounin (1970).

Rewards and punishment. Overall, the association between punishment and outcome was weak and inconsistent. The relationship between rewards (praise, appreciation) and outcome was more consistent, and all forms of reward tended to be associated with better outcomes. Note, however, that teachers infrequently used rewards (three or so instances of praise per lesson on average).

BEST COPY AVAILABLE

Conditions of learning. Independent of specific rewards, schools varied in the extent to which they provided a pleasant, comfortable environment for students. One might expect good working conditions to encourage pupils to appreciate the school and perhaps to identify with its goals. To investigate this possibility, Rutter et al. (1979) developed a 14-item scale measuring general school conditions such as freedom to use the building during breaks and the lunch period, access to the telephone, and so on. High scores on this scale were associated with higher scores on exams. As a case in point, when pupils were asked whether, if they needed to, they would talk to a member of the staff about a personal problem, a higher proportion of students in the schools with better attendance and academic achievement said that they would.

Responsibilities and participation. Pupils were asked to describe the extent to which they were encouraged to take responsibility and to help manage their school lives. The proportion of students who had some sort of school role (e.g., team captain, homework monitor, school assembly, etc.) varied greatly across schools (from 7% to 50%), and this proportion in each school correlated significantly and positively with classroom behavior and academic success.

Major Conclusions

Rutter (1983) suggests major conclusions from his earlier work include the following:

1. Secondary schools in inner London differed markedly in the behavior and achievement shown by students.

2. Although schools varied in the proportion of behaviorally difficult or low-achieving children they admitted, these differences did not completely account for variations among schools in their pupils' subsequent behavior and attainment. This provides strong evidence that school factors affected students' behavior and achievement.

3. The variation among schools on the different pupil outcome measures was reasonably stable over periods of four or five years.

4. In general, individual schools performed fairly similarly on all outcome measures. That is, schools in which students had better-than-average behavior also had students with better achievement and less delinquency. There were some exceptions in this pattern, but the trends were substantial.

5. Differences in outcomes between schools were not due to physical factors such as size of the school or the age of the building.

6. The differences among schools in outcomes were systematically related to school characteristics (e.g., identifiable factors in academic emphasis, teacher behavior, etc.).

7. Outcomes were also influenced by factors outside teachers' immediate control. For example, examination success tended to be better in schools with a substantial nucleus of children of at least average intellectual ability, and delinquency rates were higher in schools with many of the least able pupils.

8. The effect of balance in intake was most marked with respect to delinquency and least important in the case of students' classroom behavior.

9. The association between the combined measure of overall school process and each of the outcome measures was much stronger than was the relationship between any individual process variable and outcome measure. This suggests that these various social factors may combine to create a school ethos, or set of values, attitudes, and behaviors which characterize the school.

10. The total pattern of findings indicates a strong probability that the association between school process and outcome reflects at least a partially causal process.

This is a carefully conducted, important study. The data provided by Rutter et al. strongly suggest that school process has important effects on student outcome measures. Indeed, as noted earlier, Brookover et al. found that some white elementary schools obtained academic achievement scores an average of about one standard deviation above those in less successful white elementary schools matched for intake. The difference in achievement between the successful and unsuccessful black schools was somewhat greater.

Still, school inputs (more favorable student populations) were positively correlated with student outcomes and although strong and persuasive arguments

can be made that school process affected outcomes more than did input variables (e.g., see Rutter, 1983) the data are correlational. Subsequent field experiments will help to determine if processes identified by Rutter et al. as characteristic of effective schools can be manipulated in ways that lead to improved outcomes in other schools. This study collected more process data of better quality than did other school effects studies. However, the sample of teachers observed in each school and the range of process variables studied provide only a limited view of school life. The study has nevertheless provided a solid foundation upon which subsequent experimental and observational studies can build.

Integrative Reviews of School Effects Research

Not only have there been many attempts to define and study effective schools, there are also several literature reviews order and integrate the studies that have been conducted. Purkey and Smith's (1983) review is one of the most comprehensive. Their review is especially instructive because it includes a wide range of research approaches representing different theoretical positions.

Purkey and Smith (1983)

These researchers argue that it is easy to conclude that the findings of recent school effects research contradict the conclusions of Coleman et al., Jencks et al., and others. However, they note that new studies do *not* refute the earlier findings that easily measured structural differences among schools (library resources, etc.) are not consistently related to student achievement. Further, although recent research does not indicate that there are large differences in achievement among schools, this research has important implications for practice.

Purkey and Smith's review examines various types of school effectiveness research including outlier studies, case studies, surveys and evaluations, as well as studies of program implementation and theories of organization in schools and other institutions.

Outlier studies. This type of school effectiveness research statistically identifies highly effective and unusually ineffective schools and then examines behavior in these schools to determine what accounts for the differences. Studies using the outlier approach included in the review were four studies conducted by the New York State Department of Education (1974a, 1974b, 1976); a study conducted for the Maryland State Department of Education (Austin, 1978); as well as research by Lezotte, Edmonds, and Ratner (1974); Brookover and Schneider (1975); and a study of Delaware schools by Spartz, Valdes, McCormick, Myers, and Geppert (1977).

These studies generally show that the most common elements of effective schools are better control or discipline and high staff expectations for student achievement. Each of these variables was evidenced in four of the seven studies for which there are data. An emphasis on instructional leadership by the principal or another important staff member was found to be important in three studies.

Purkey and Smith note that these outlier studies are similar in some respects; however, variations in the findings from this research should serve as a caution to those who would reduce the school effects literature to five or six variables.

Case studies. Eight school case studies were reviewed: Brookover et al., 1979; Brookover and Lezotte, 1979; California State Department of Education, 1980; Glenn, 1981; Levine and Stark, 1981; Rutter et al., 1979;

Venezky and Winfield, 1979; and Weber, 1971. The inherent weakness of the case study approach is the small sample size; however, the commonality of findings among these studies and the similarity of their results to findings from other kinds of studies increase their credibility.

After reviewing six of the case studies that examined a total of 43 schools (Rutter et al. and Brookover et al. were discussed separately because of their greater complexity), Purkey and Smith note that each case study focused on urban elementary schools and that the studies varied in quality of methodology and clarity of reporting.

Five factors were common to most, but not all, of the six case studies in this group: (a) strong leadership by the principal or other staff, (b) high expectations by staff for student achievement, (c) clear goals, (d) an academic emphasis for the school and an effective school-wide staff training program, and (e) a system for monitoring student progress. A focus on order and discipline was found to be important in two of the studies and a large number of factors were specific to any single study.

Program evaluations. Purkey and Smith also examined six program evaluation studies: Armor et al., 1976; Doss and Holley, 1982; Hunter, 1979 (three studies carried out by the Michigan Department of Education); and Trisman, Waller, and Wilder, 1976. They note that though these studies are methodologically stronger than the preceding two types of research, their findings are remarkably consistent with the other studies.

Portrait of an effective school. Purkey and Smith found the following variables to be important process measures of school effectiveness: (1) school-site management--a number of studies indicate that leadership and

staff of the school need considerable autonomy in determining how they address problems; (2) instructional leadership--though the reviewers are suspicious of the "great principal" theory, it seems clear that leadership is necessary to initiate and maintain school improvement; (3) staff stability--once a school experiences success, retaining the staff seems to maintain effectiveness or to promote further success; (4) curriculum articulation and organization--at the secondary level, a planned, purposeful program of courses seems to be academically more beneficial than an approach that offers many electives and few requirements; (5) school-wide staff development--essential change involves altering people's attitudes and behavior as well as providing them with new skills and techniques; (6) parental involvement and support--though the evidence on this issue is mixed, it is reasonable to assume that parents need to be informed of school goals and school responsibilities; (7) school-wide recognition of academic success--the school culture is partially reflected in the ceremonies, its symbols, and the accomplishments it officially recognizes; (8) maximized learning time--if schools emphasize academics, then a greater proportion of the school day will be devoted to academic subjects; and (9) district support--fundamental change, building-level management, staff stability, and so on all depend on support from the district office.

Purkey and Smith believe that other process variables must be present as well: (a) collaborative planning and collegial relationships, (b) sense of community, (c) clear goals and high expectations, and (d) order and discipline.

Other Synthesis Reviews

Cohen. Cohen (1983) provides an overview and framework for interpreting the rapidly accumulating knowledge base concerning schooling practices that contribute to student achievement. He notes that existing summaries are fine

as far as they go, but that by presenting only lists of variables, they fail to provide information on how these factors are interrelated, how they can actually be implemented, or their effects.

Cohen points out that by attempting to explain differences between schools' average level of student achievement, most previous research overlooks the important fact that most of the variance in student achievement (between 70% and 90%) actually occurs *within* schools. Also, the exclusive focus on average differences between schools assumes that all school resources are equally available to and utilized by each student in a school. Yet, within schools many students are grouped into tracks and within classes into ability groups. These groups are exposed to different teacher skills, instructional practices, curriculum materials, and social environments, all of which are believed to influence school learning.

Cohen suggests three characteristics of effective schools that can be used to organize existing research. First, school effectiveness is clearly dependent upon effective classroom teaching. Second, school effectiveness requires the careful coordination and management of the instructional program at the building level. Finally, effective schools generate a sense of shared values and culture among both students and staff.

Cohen notes that research on school practices is *not* as well developed as that on classroom practices; there are fewer studies, less frequent replication of findings across studies, and fewer concrete descriptions of specific behaviors and practices. Despite these deficiencies there are several important general research findings concerning the management and coordination of instruction. First, the curriculum and instructional programs in effective schools, especially elementary schools, are interrelated. This means that school goals, grade-level and classroom instructional objectives,

instructional content and activities, and measures of pupil performance are all carefully coordinated such that the instructional efforts of teachers and other instructional staff are consistent and additive.

This interconnectedness among several elements of the instructional program has several implications. First, it requires that schools have clear, public, and agreed-upon instructional goals that form the basis for selecting objectives, content, and materials. Second, the interrelationships imply that extreme differences in time allocation to the same content do not exist within a school. Cohen argues that extreme time allocation differences probably reflect the substitution of teacher preferences for the formal academic goals of the school, and result in students in various classes being exposed to functionally different curricula (see Berliner, 1979). Interrelatedness also implies that prevailing norms which grant considerable autonomy to teachers behind the closed door of the classroom carry less weight than do the shared goals of the professional staff. Furthermore, the expectations and instructional activities of nonclassroom specialists (e.g., resource teacher, reading specialists, etc.) are consistent with and supportive of the efforts of the classroom teachers. Finally, there is overlap between the content of instruction and the content included in measures of pupil performance. If textbooks and tests are not carefully coordinated (either at the school or district level), test results do not accurately reflect student learning. For an extended discussion of this problem, see Freeman et al. (1983).

According to Cohen, this argument for coordination of curriculum and instruction is *obvious*, just as a description of effective teaching practices seems *obvious*. However, coordination in schools apparently occurs infrequently. Furthermore, coordination of goals and content with performance measures

suggests an image of schools that conforms closely to a classical model of bureaucracy. However, he notes that descriptions of the organization of effective schools differ notably from what is believed about schools in general. He suggests that perhaps effective schools are different from most other schools; in particular, they are better managed, their work is more frequently directed toward appropriately limited, shared goals; and instructional practices are more advanced and consistent with the most recent research.

Edmonds. Ronald Edmonds, until his untimely death in 1983, had been one of the key figures in the school effectiveness movement. His work represented a major integrative attempt to demonstrate that schools are not interchangeable and that some schools have much more impact than others with similar resources serving similar populations. Edmonds, more than anyone, had been responsible for the communication of the belief that *schools* can and do make a difference.

In addition to the basic issue of school effectiveness (Do schools make a difference?), Edmonds spent much time dealing with (a) the investigation of pupil background characteristics that affect school effectiveness, (b) the generality of school effectiveness, (c) comparison of methods for evaluating school effectiveness, (d) the equity of proposed evaluation procedures, and (e) the comparison of effective and ineffective schools (Edmonds, 1983).

His most salient contribution was the articulation of a model for characterizing effective schools. His model underlies many of the school improvement models now being implemented in many American schools. Edmonds (1982) contended that

> the correlates of effective schools are (a) the leadership of the principal characterized by substantial attention to the quality of instruction, (b) a pervasive and broadly understood instructional focus, (c) an orderly, safe climate conducive to teaching and learning, (d) teacher behaviors that convey the expectation that all students are to obtain at least minimum mastery, and (e) the use of measures of pupil achievement on the basis for program evaluation.

Two of the many school improvement projects that have been stimulated in part by Edmonds' work follow.

Application of Findings from School Effects Research

At present there is clear evidence that many schools, school districts, and state departments of education are applying the results of school effectiveness research in order to improve student performance. Nevertheless, there is a paucity of data that indicate how well or in what way research findings are implemented in particular schools and how various forms and processes of implementation relate to student performance. Fortunately, however, many evaluation projects are in progress and three of these efforts will be discussed in this section.

Milwaukee's Project RISE

McCormack-Larkin and Kritek. McCormack-Larkin and Kritek (1983) note that although the quest to identify the characteristics of instructionally effective schools continues, some urban school districts are currently using available evidence to design and implement programs. One such program began in Milwaukee in March 1979 when the school board directed the administration to develop a plan for improving achievement in 18 elementary and two middle schools that scored lowest on annual achievement tests. Project RISE is based upon three assumptions. First, virtually all students, regardless of their family background, race, or socio-economic status, can acquire the basic skills. Second, inappropriate school expectations, norms, and practices account for the underachievement of many low-income and minority students. Third, the literature on effective schools and classrooms identifies expectations, norms, practices, and policies that are associated with high achievement.

In particular, the researchers believed that eight features needed to be implemented in school programs: (a) the belief among faculty and students that all students can learn and that the school is primarily responsible for their learning; (b) a strong sense of academic mission; (c) a high level of professional collegiality among staff members; (d) a strong sense of student identification and affiliation with the school; (e) grade level expectations and standards in reading, math, and language; (f) an accelerated learning program for students performing well below grade level (a component of the accelerated learning program is whole-class instruction at grade level, supplemented by small-group instruction at the students' actual skill levels); (g) increased time used for active student learning; and (h) a structured learning environment.

RISE schools abandoned the system-wide ungraded, continuous-progress school organization approach and developed grade-level objectives and standards. Grade-level standards were the prerequisites for success at the next grade level. Teachers used these standards as a guide in their planning, and students and parents were informed of them by a checklist.

Project RISE completed three years at the end of the 1981-1982 school year. Although the program continues, it is reasonable to ask now whether or not it has achieved its goals. Results on a standardized test administered by the school district appear in Table 3.

McCormack and Kritek note that RISE schools have shown improvement in mathematics and some improvement--though not as dramatic--in reading. They note that the percentage of students in Project RISE who scored in the average or high range on the MAT has increased by about 25% in mathematics, and mathematics performance of Project RISE students is now comparable with that of other students in the district. Differences in achievement in the two subjects can possibly be attributed to the comprehensive instructional inservice

Table 3

Percent of Students in Milwaukee in Average or High Categories on MAT

School Year	City Avg.	RISE	Difference
Third Grade Reading			
1975–1976	73	51	22
1976–1977	68	47	21
1977–1978	70	53	17
1978–1979	74	53	21
1979–1980	76	59	17
1980–1981	77	64	13
Second Grade Reading			
1981–1982	68	64	4
Third Grade Math			
1975–1976	73	56	71
1976–1977	68	52	16
1977–1978	71	52	19
1978–1979	75	56	19
1979–1980	81	73	8
1980–1981	84	82	2
Second Grade Math			
1981–1982	80	80	0
Fifth Grade Reading			
1975–1976	65	36	29
1976–1977	64	38	26
1977–1978	65	39	26
1978–1979	66	45	21
1979–1980	68	48	20
1980–1981	72	56	16
1981–1982	71	58	13
Fifth Grade Math			
1975–1976	66	40	26
1976–1977	64	44	20
1977–1978	68	50	18
1978–1979	70	58	12
1979–1980	73	66	7
1980–1981	80	77	3
1981–1982	83	79	4

programs provided by a committed group of Title I math teachers. These teachers worked with classroom teachers, using the key instructional behaviors included in the eight features outlined above.

Still, the findings raise some areas of concern. Rutter et al. (1979) and Rutter (1983) argue an "ethos effect," yet when schools have uneven effects across different outcome measures and for various students, it is difficult to understand why a climate variable like ethos should have diverse

outcomes in various subject areas. Clearly, presumed *general* features of effective schools cannot explain effects that are specific to subject matter.

Milwaukee school board policy has changed on two important issues, at least partly because of the influence of Project RISE. First, the schools have shifted from an ungraded primary to a graded system. As of the 1983 school year, Grades 1, 2, and 3 exist officially for the first time in approximately 40 years, and grade-level objectives have been adopted system wide. Second, schools can now retain primary students who have not mastered the skills deemed essential for the next grade level. It is possible that the retention of students in a grade may partially explain higher scores in future testing. (If a greater percentage of students repeat a grade level, part of the gain may be due to additional instruction.)

The investigators note that Project RISE has required only a small amount of money, and no new staff, new materials, and so on. Nevertheless, Project RISE appears to have achieved some success. For example, it is clear that one concrete goal of the program--that of raising district achievement--has been achieved to some extent, especially in some schools and in the area of mathematics.

New York City's School Improvement Project

Clark and McCarthy. Clark and McCarthy (1983) note that the School Improvement Project (SIP) is based on five factors derived from the school effectiveness literature, which focuses on causes of school failure (Edmonds, 1979). These five factors are administrative style, instructional emphasis on basic skills, school climate, ongoing assessment of pupil progress, and teacher expectations. SIP builds on typical school effectiveness research because it sets goals for school improvement and establishes a system for obtaining these goals. This project is a response to one criticism of the school effectiveness literature, that longitudinal studies of improvement projects in a variety of schools have not been conducted.

BEST COPY AVAILABLE

This school change project is also based on Edmonds' (1979) contention that schools can educate all children, regardless of family background. In June, 1979 the Chancellor of the New York City public school system issued a memorandum to community school districts and all principals in elementary schools, inviting them to submit applications for participation in New York City's School Improvement Project. Of the 43 schools that applied, 10 were selected to participate in 1979-1980. Three criteria were used to select schools: voluntary participation of the principal (with the superintendent's approval), a match between school needs and SIP objectives, and a lack of other development programs in the school. An additional nine schools from the original pool joined the project in 1980-1981. In the third year, four new schools entered SIP at the discretion of community superintendents.

Most students in eight of the 10 schools in the first cohort were minority, and over half were designated as low income. From 20% to 61% of the students in a school were at or above grade level. During the first project year, SIP schools conducted a needs assessment, formed school planning committees, developed improvement plans, and reviewed and refined the plans. The plans were implemented during the second year, and strengths and weaknesses were identified.

Implementation of revised plan activities in first-cohort schools was considered somewhat successful by principals and liaisons. The most successful plan activities implemented in all seven schools were reading and language arts programs and test sophistication programs in reading and mathematics. The original schools implemented an average of 14 plan components, in comparison to 21 components instituted by second-cohort schools in 1981-1982.

Preliminary student achievement data from spring 1979 to spring 1982 (see Table 4) indicate the percentage of students reading at and above grade level

BEST COPY AVAILABLE

Table 4

Student Reading Achievement in SIP Schools and City-wide: Percent At and Above Grade Level on the California Achievement Test, Grades 2 to 6, Spring 1979 to Spring 1982[a]

	Percent and Number At and Above Grade Level[b]							
	April 1979		April 1980		April 1981		April 1982	
	%	N	%	N	%	N	%	N
First-Cohort (7)	26.0	869	33.7	1,061	40.6	1,220	41.0	1,178
Gains			+7.7%		+6.9%		+0.4%	
Second-Cohort (8)[c]	—	—	33.5	1,306	39.3	1,434	44.5	1,487
Gains			—		+5.8%		+5.3%	
Combined-Cohort (15)[d]	—	—	33.6	2,367	39.9	2,654	42.8	2,665
Gains			—		+6.3%		+2.9%	
City-wide-Cohort (690)[e]	39.6	124,986	47.5	143,127	51.3	148,581	51.0	142,468
Gains			+7.9%		+3.8%		-0.3%	
Difference Between SIP and City-wide			-0.2%		+2.5%		+3.2%	

Note. From "School Improvement in New York City: The Evolution of a Project" by T.A. Clark and D.P. McCarthy. 1983. Educational Researcher, 12(4), p.22.

[a]Since all SIP schools are either K-5, or K-6, city-wide scores were used only through sixth grade. New York City students in Grades 2 through 9 are given the C annually. Many special education and certain limited-English proficient students are excused.

[b]Because enrollments fluctuated from year to year in every school, percentages are based on actual enrollment figures for each year rather than on the same enrollment figures.

[c]The second-cohort schools started in 1980-1981, so data are not provided for 1979.

[d]Although 19 schools entered SIP in 1979-1980 and 9 in 1980-1981 schools dropped out during or right after the first year. Only 15 schools which completed at least 2 years are included in the combined cohort.

[e]SIP schools are not included in city-wide figures. Included are 609 elementary schools and Grades 5 and 6 in 81 intermediate schools.

each year as measured by the California Achievement Test. For two of the three years data were analyzed, SIP schools showed greater increases than other city schools in percentage of students reading at or above grade level. For example, between spring 1980 and spring 1981, 15 SIP schools had a 6.3% increase, as compared to 3.8% for city-wide schools. Between 1981 and 1982, SIP schools continued to show increases, while city-wide schools decreased by .3%.

According to Eubanks and Levine (1983), achievement data in the New York City Project are encouraging. They note that the seven schools that initiated improvement plans in 1980-1981 had an average increase of 16 points between spring 1979 and spring 1982 in the percentage of students reading at or above grade level. However, they also note that these achievement gains have stabilized and that only a few third-year schools registered noticeable gains

in 1981-1982. It would be important to know more about the plans and their degree of implementation in schools where improvement continues.

The project administrators believe that vital leadership came from principals; in schools where the plan worked, the principals supervised and coordinated implementation of the plan components and monitored results closely.

One implication of this project is that researchers need to obtain teachers' and principals' commitments to the SIP program. Indeed, the project now requires that principals, 60% to 70% of the staff, and the parents agree to participate through a formal vote. The program currently emphasizes that the liaison acts as a facilitator and the principal as a supervisor of the process.

Other School Improvement Efforts

The general programs discussed in the above section are but a *few* of the many school improvement projects under way. Despite many problems associated with the extant data base available for designing school improvement plans, many projects are in progress. Unfortunately, some districts have simply taken plans developed in other districts and applied them with few, if any, modifications. However, persons who have studied the literature rightfully advocate that schools or school districts need to develop plans that are relevant to their unique needs and populatio n. Furthermore, districts must understand the change *process* if schoo. effectiveness literature is to be used appropriately.

Purkey and Smith (1985) aptly summarize the implications of recent research for schools.

> It would mean, for schools, increased involvement of teachers and other staff members in decision making, expanded opportunities for collaborative planning, and flexible change strategies that can reflect the unique 'personality' of each school.

BEST COPY AVAILABLE

Thus, unless districts utilize appropriate change strategies, pressure for immediate school change will lead to the attempt to improve schools through uniform policy standards (e.g., all schools will have homework, school plans, three years of mathematics, etc.) and may negatively affect the achievement of pupils in schools that are already most effective. Finn (1983) reasons that effective schools have become so because they have developed their own goals, norms, and expectations. He suggests that the problem is not that all schools will be required to have higher standards per se, but that the mandate to have such standards may erode the pride and the continuing commitment of good schools to seek better solutions to educational problems.

Finn (1983) argues this point:

> The schools in a given system or state are apt to be similar with respect to relatively superficial matters but dissimilar along dimensions that matter more; yet the inertial autonomy of schools *qua* schools also means that efforts to make ineffective schools more closely resemble effective schools in the ways that matter most are certain to be very difficult and quite likely to meet with little success. Moreover, policymakers seeking greater uniformity must be terribly careful lest they "level downward" through well-intentioned efforts that wind up sapping the vitality of the most effective schools rather than invigorating the others. (p. 5-6)

Goodlad (1983) provides empirical support for Finn's contention that schools appear the same on many superficial dimensions but vary in important, subtle ways. He suggests that it is futile to attempt to improve many ineffective schools unless these schools demonstrate some basic initiative towards change. Goodlad writes,

> But the less satisfying schools described are not sufficiently in charge of their own destiny to build agendas from these data and subsequently to take action. They are not healthy organisms. They simply are not good candidates for tackling the difficult tasks of curricula and pedagogical reform. The first step is for them to become more effective in the conduct of present business and in the process to become more satisfying places for those

> associated with them. They need data about their present condition and considerable support and encouragement in the needed process of renewal. (p. 269-270).

Thus, if schools do not have agreement among staff on school-wide goals, telling them to develop such goals may be irrelevant. Such schools may be able to change only after they are convinced that change is necessary, receive outside help in assessing curriculum and instruction, and in some cases, specific information about how to bring about needed change.

Many school improvement efforts begin with a needs assessment survey that solicits input from the school staff to outline school problems and general needs. However, one could argue that school staff may not be able to identify certain problems and issues, and that external review might help schools to analyze their current strengths and weaknesses. For example, the descriptions of schools provided by Jackson (1981a, 1981b) and by Lightfoot (1981a, 1981b) not only differ from each other but would likely differ from descriptions provided by the staffs of these schools. There are many of views what a school is and should be, and both internal and external assessments may be necessary in order to describe accurately a school's problems and potential.

Criteria for Evaluating School Plans

Although there has been considerable talk about the need for school plans, researchers and educators rarely describe individual school plans in journal articles or technical reports. By what criteria can we judge school plans? Implicitly, the criteria for judging plans seem to be the percentage of faculty involved in developing a plan and the number of faculty who accept the plan as valid. There can be bad plans as well as good ones, however, and some plans might constrain experimentation and growth, while others would encourage these activities. On the other hand, a school must start somewhere. Some schools will never change unless external pressure is brought on them to do so

and although staff agreement is desirable, the best that can be done in some situations is to work around those who are uncooperative but entrenched.

Data from studies of school improvement are less likely to explain *why* results are obtained or the most promising ways to develop school improvement plans. Although the plans are said to be stimulated by school effects and teacher effects research, it is often difficult to identify their research basis. In general, documentation focuses on the extent to which teachers, parents, and administrators are involved in developing the plans, and far too little attention is paid to the *validity* of the plans themselves or the extent to which they are implemented in actual practice.

Implementation of School Plans

Information is especially needed concerning the extent and quality of school plan implementation in individual classrooms. Without descriptions of *intended* and *actual* implementation of instructional, organizational, and social processes, it is difficult to assess why achievement increases in some schools and not in others. The seriousness of this question is reflected in the observation of Purkey and Smith (1983a) that only one-half of the schools participating in Milwaukee's Project RISE showed achievement increases. Were achievement increases due to motivational factors or to identifiable changes in instructional behavior?

Other school improvement projects report similar variation. For example, McCarthy, Canner, and Pershing (1983) note that 60% of the schools in the New York City Project were successful. They further n te that schools implemented instructional plans more successfully than plans related to improving school climate, parental participation, or school administration.

Future School Evaluation Plans

Future school evaluations must have better conceptual statements of what the school improvement program is (What is the independent variable?), improved statements of theory (Why is the program expected to work?), and more thorough observational evaluations (preferably by persons who are independent of the school district) of what takes place in schools. Too much of the current school improvement activity naively proceeds as though the existence of a school plan that is widely accepted will positively influence achievement outcomes.

The Process of Implementation

Fullan (1985) extensively examines how change processes work. He contends that an awareness of the conditions of change is critical if the effective schools literature is to be appropriately applied. His paper discusses a number of issues related to school change--Should participation be mandatory? At what level should the process be implemented (individual classroom, individual school, or district level)? How much variation should be encouraged (at the classroom level, at the building level)? Should large- or small-scale approaches be used?

He contends that the process of implementation is crucial. Although it is beyond the scope of this paper to discuss the change literature, it is useful to highlight briefly the steps that Fullan recommends in helping participants to develop new skills in a manner that is satisfying and meaningful to participants:

1. Change takes place over time;
2. the initial stages of any significant change *always* involve anxiety and uncertainty;
3. ongoing technical assistance and psychological support are crucial if the anxiety is to be coped with;

4. change involves learning new skills through practice and feedback--it is incremental and developmental;

5. the most fundamental breakthrough occurs when people can understand the underlying conception and rationale with respect to "why this new way works better;"

6. organizational conditions within the school (peer norms, administrative leadership) and in relation to the school (e.g., external administrative support and technical help) are important;

7. successful change involves pressure, but it is pressure through *interaction* with peers and other technical and administrative leaders.

Limitations of Extant Research

Although the literature yields a number of statements concerning the characteristics of effective schools, there are several limitations that must be considered when one uses extant findings for program planning. These limitations are similar in many ways to the problems of implementing findings from teacher effects research (e.g., Berliner, 1977; Good, Biddle, & Brophy, 1983); however, the issues related to school effectiveness are more complex because there are fewer observational data that describe school processes, and independent investigators have not conducted carefully controlled field experiments.

Independent Variables

To date not a single naturalistic study of effective schools provides basic data (means and standard deviations for each classroom) to demonstrate that the behavior of individual teachers in one school differs from the behavior of teachers in other schools. It will be important for future research to examine how schools vary in terms of various quantitative measures (e.g., How much homework is assigned per class?) and qualitative measures (e.g., Does homework consist largely of drill or is it designed to develop independent study skills? How do teachers use such work--do they examine it, comment on it, and use it as part of the basis for assigning a grade?).

Validity

In this paper we have described studies that allow one to determine if and how schools influence average student performance. Although information about students' performance on criterion-referenced and norm-referenced achievement tests is valuable (if those tests are congruent with the curriculum actually used in the classroom), academic achievement is only part of what many citizens and educators mean by an effective teacher or an effective school. Most studies have examined only student achievement (studies by Brookover et al., 1979 and Rutter et al., 1979 are important and notable exceptions).

As a case in point, Miller (1983) argues and provides data to support his contention that the scientific literacy--the ability to read about, understand, and express an opinion on scientific issues--of most Americans is appallingly low. Arguments could also be made about students' performance in other areas. For example, Goodlad (1983) contends that students are very passive in their reaction to school events and that schools should help to stimulate more student thinking. Hence, there are many possible outcomes of schooling, and although it is important to acknowledge and to utilize appropriately new data about effective schools, it is inadvisable to equate school effects on achievement with effectiveness per se, because many important outcomes of schooling have not been examined. (Others, too, have reached similar conclusions; see Cuban, 1983; Purkey & Smith, 1983a; Ralph & Fennessey, 1983; and Rowan et al., 1983).

Stability

It would be disturbing to go to the trouble and expense of identifying effective schools only to learn that these schools do not have stable effects on student achievement across consecutive years. This would be the equivalent

of attempting to determine what makes an effective football team by identifying a football team that is outstanding during one year (e.g., the 1982 World Champion San Francisco Forty Niners) and studying that team intensively the following year (e.g., policies related to scouting, drafting, training, game strategy, etc.), only to discover that the team did not even make the playoffs during the year when observational measures were collected.

Available evidence does not provide generalizable information about the stability of effective schools. It does appear that the *average* stability of school effects on student achievement is low. Rowan and Denk (1982) estimate that only about 10% of the schools drawn from a large sample were consistently effective or ineffective (i.e., in the top or bottom quartile of the residual distribution over consecutive years), and only 5% were effective or ineffective over three consecutive years.

Rowan and Denk (1982) used data provided by the California Assessment Program (California State Department of Education, 1977) to construct two measures of instructional effectiveness for sixth-grade achievement in 810 classes from 1975-1977. The two measures they used were (a) trend analysis (Is the school improving over consecutive years?) and (b) regression-based residuals. They note that the trend analysis was biased against schools serving low-SES populations and that decreases in scores from year to year were significantly correlated with changes in the socio-economic composition of these schools' student bodies. Moreover, they found that this measure of effectiveness lacked stability; the correlation between gains made from 1975 to 1976 and gains from 1976 to 1977 was -.45.

These researchers also examined the bias and stability of a measure based on residuals. As expected, this measure was not correlated to measures of student background; however, the measure was relatively unstable. For

example, the correlation of residuals for 1975 and 1976 was only .24; for 1976 and .19 for 1977. These results are consistent with other studies of the stability of residuals (e.g., Forsythe, 1973 and Jencks et al., 1972). Thus, the stability of school effectiveness measures varies according to the procedure used to estimate stability; however, all comparisions indicate that the average stability of school effects on student achievement is low.

Some data suggest that performance in "effective" schools may be more stable than in other schools. Recall that in Weber's (1971) study two of the schools (50%) maintained their effectiveness over consecutive years. Rutter et al. (1979) and Rutter (1983) report that schools that were effective (achievement, attendance, behavior) generally maintained their effectiveness over several consecutive years.

Other evidence also suggests that at least some aspects of school effects are stable over time. Reynolds, Jones, and St. Leger (1976) found a correlation of .85 for school attendance rates over seven years and of .56 for academic attainment over the same period of time. Such stability justifies the study of effective schools and attempts to describe successful practice in these schools. However, the study of stability presents major technical and conceptual problems to those who study schools as organizational, instructional units. Why do some schools achieve highly one year but not the next? If strong principal leadership is an important variable in school achievement, how and why does achievement vary from year to year?

Multiple Criteria

In addition to questions of the stability of school effects on students' performance, there is also the issue of whether schools are generally effective or whether effectiveness is limited to a few areas of student performance. Unfortunately, this important question has received little attention;

most studies of school effectiveness include only achievement measures. The data that do exist suggest that effective schools can positively influence different outcome measures simultaneously. For example, Rutter et al. (1979) found that schools that had favorable effects on student achievement also generally had positive effects on attendance and classroom behavior. (See also related evidence reviewed by Rutter, 1983).

However, to repeat, there is little evidence directly related to the question of how general school effects are, and the outcome measures used to assess generality appear to be logically related to achievement (attendance, classroom behavior, engagement rates, etc.). Whether schools that help students to learn to express themselves well through written essays, to develop adequate computer skills, and/or to understand scientific processes well are the same schools that have high average scores on standardized achievement tests is an open question that needs to be explored.

Practical and Conceptual Issues

It seems unreasonable to include data from pupils who do not attend a school in analyses measuring school effectiveness. Despite the obvious nature of this observation, many schools with high turnover rates include in their achievement data scores of numerous students who received only limited instruction in the school. If valid and reliable measures of school effects are to be obtained, investigators must establish and report minimum instructional time periods for including students or teachers in a data set. One reason the stability of residual gain achievement scores is low in some studies may be that not enough care has been taken to determine whether a student has been "taught" in a particular school and to associate the student *with particular teachers and classes*. For example, Averch et al. (1974) note that in their extensive review of input-output studies only one study

(Hanushek, 1970) matched student achievement with resources that teachers and students were actually exposed to.

Rowan et al. note that most studies of effective schools do not measure the instructional performance of an *entire* school. Rather, schools are labeled as effective on the basis of assessments of instructional outcomes at only one or two grade levels and in only one or two curriculum areas (for an exception, see Wellisch, MacQueen, Carriera, & Duck, 1978). Also, these authors argue that even within curriculum areas and at a single grade level, schools may not be uniformly effective for all types of students. Marco (1974) studied 70 Title I schools and found a correlation of only .32 between indexes of instructional effectiveness calculated separately for students of high and low ability (cf. Shoemaker, 1982).

Context

Only a few types of schools have been studied extensively. For example, researchers have more frequently attempted to identify effective inner-city schools and rural schools than to study effective suburban schools. Also, most attempts to improve schools have focused on elementary schools. The implications of the effective schools literature for suburban or secondary schools are therefore uncertain.

Correlational Evidence

Most of the research attempting to associate school effects with student learning is correlational. Although this information is valuable, correlational evidence only allows one to infer that two variables are associated, not that either variable directly influences the other. As Rowan et al. (1983) point out, the finding that strong principal leadership is associated with clarity of instruction and student achievement does not necessarily mean

that effective principals cause teachers in their schools to be focused and systematic. It may well be that a capable teaching faculty preceded leadership and enabled the principal to assume a strong and active role.

Furthermore, correlational statements describe a relationship at one point in time. If one assumes that existing literature is valid and that certain factors are associated with effectiveness (e.g., leadership), research only indicates that principal leadership and student achievement are correlated; it does not provide evidence about how one can become a leader.

Nonlinear Relationships

Many teacher effects researchers argue that most teacher behaviors have a nonlinear relationship with student achievement (Berliner, 1977; Good, Biddle, & Brophy, 1975; and Soar & Soar, 1979). A linear relationship means that a teacher behavior and a student outcome measure are directly related--increases in the teacher behavior are accompanied by increases in student achievement. Research on teaching has often found nonlinear relationships between classroom processes and student achievement. One relatively common nonlinear pattern is an "inverted-U." In this pattern, outcome measures are associated with an optimal level of a classroom process (or teacher behavior), so that teachers who exhibit too little or too much of this process have less positive effects on students than do teachers who use the correct amount.

Many school effects variables probably have nonlinear relationships with outcomes as well, but the paucity of observational data collected in this research makes it impossible to assess the claim. Still, it seems likely that a school could have either too few or too many rules, or could assign either too much or too little homework. Even teacher expectations can be too high as well as too low (Edmonds, 1983; Good & Brophy, 1984). Unfortunately, too many

advocates of change based on the school effectiveness literature assume linear relationships--if a little is good, more is certain to be better.

Averages Can Be Misleading

Studies of large samples of schools yield important profiles of more and less successful schools, but these are usually *group averages* that may or may not describe how a single effective teacher actually behaves in a particular effective school. Persons who use research to guide practice sometimes expect all teachers' behavior to reflect the group average. Such simplistic thinking is apt to cause the literature to be too broadly and inappropriately applied.

It is also important to consider the *appropriateness* and *effects* of schooling for specific types of students. It would seem important to make mathematics (and other subjects) as interesting for girls as for boys, for example, and for minority as well as majority students (extant data on this point are not encouraging). It seems unreasonable to call a school excellent if one group of students is not making progress; however, if school effects are reported only in total averages, it will not be possible to discover whether certain groups experience poor growth in an otherwise good school.

School Plans/Teacher Autonomy

There is presently great interest in individual schools conducting needs assessments and developing their own plans. Many advocates of school improvement argue that because of unique student and teacher populations, community characteristics, histories, and resources, schools must have autonomy in developing plans. However, this same logic dictates that teachers and departments also need a degree of autonomy to chart their own directions and to record their progress. Ironically, many of those who argue most strongly for school autonomy are least interested in teacher autonomy. Although we believe

that some consensus on schools goals and use of resources across an entire staff is necessary, there can be too much as well as too little school-wide planning.

Whether departments in secondary schools are in agreement about curricula and instructional practices in particular subjects is also an important issue. Similarly, researchers should ask if teachers have a firm sense of the improvements they are trying to effect in their classrooms. The school improvement model that many researchers and theorists advocate regards the school as an institution that establishes goals and the mechanisms that make it likely that these goals will be realized. Another way to consider school effects, however, is as processes that encourage teachers to carefully consider what they are trying to achieve in their classrooms, to become more aware of their progress on such goals, and to make better use of other school personnel in developing their classroom strategies.

At any rate, naturalistic research does not indicate that more effective schools have more formal school plans (conduct needs assessments, etc.), but only that there is a greater informal consensus among the staff in these schools that students can learn, etc. As a strategy for building more school-wide direction and consensus, many school improvement projects emphasize the development of school-wide plans.

Lightfoot (1983) expresses a similsr concern in her book, *The Good High School*. From her intensive study of six diverse but presumably good high schools, she contends,

> In all of these schools, therefore, teachers are seen as the central actors in the educational process. Their satisfaction is critical to the tone and smooth functioning of the school. Their nurturance is critical to the nurturance of students. Each school interprets teacher rewards differently, but all of them search for a balance between the expression of teacher autonomy, initiative, and adulthood on the one hand, and the requirements

BEST COPY AVAILABLE

> of conformity, discipline, and commitments to school life on the other. (p. 341)

Sizer (1984) also addresses this point in his book, *Horace's Compromise*, which is based on his major study of American high schools. After concluding that good teachers are the critical element of successful schools, he suggests that effective schools must allow teachers autonomy but also must have standards and accountability. Hence, the issue of balance between schools' needs and teachers' needs in serving students must be addressed. Unfortunately, overapplication of the school effectiveness literature may make schools less attractive to more talented teachers.

It may be that some attention to school-wide planning and a general awareness of school goals, coupled with a focus on departmental or individual teacher plans, will be the strongest and most durable strategy for enhancing school productivity. Research and development activities need to focus on more differentiated improvement plans and plans that encourage both the coordination of school programs and initiative from individual teachers (e.g., Good & Brophy, 1984). Shulman (1983) also raises concerns about teacher autonomy in the context of applying school effectiveness research.

Who is Responsible for Achievement?

Most persons who write about effective schools focus on teachers and principals. However, some authors suggest that the role of the student has been modified so that less is expected of students today than in the past. Further, these persons believe that too much attention is paid to the teacher, and that students need to be held accountable for learning.

Tomlinson (1981) argues that although socially imposed standards of conduct were strict and pay was meager, teachers in the 1950s were personally consoled by the public's esteem and gratitude and by their own generally

unquestioned and unchallenged classroom authority. Teachers, especially as they grew older, might be viewed as a little eccentric, but seldom as ineffective. Their job was to expose students to information and to inculcate knowledge. Still, whatever they taught and however they taught it, they were rarely held responsible for how much a student learned. The responsibility for acquiring knowledge, or achievement, lay with the students themselves. It was their variable ability and motivation that determined how much they learned. Tomlinson notes that things have changed in recent times. The ability and effort of teachers have replaced the ability and labor of students as the putative determinants of achievement.

Tomlinson suggests that research generally indicates that academically effective schools are merely schools organized to pursue learning consistently. Principals, teachers, students, and parents agree upon the purpose, justification, and methods of schooling. They systematically spend their common energies on teaching and learning. They are serious about, even dedicated to, the proposition that children can and shall learn in school. No special treatment and no magic are involved, just provision of the necessary conditions for learning.

Other writers stress the importance of the home and general community in supporting learning activities (see for example, Lightfoot, 1978). Schools (and teachers) are important but not exclusive factors in facilitating students' learning. Unfortunately, many accounts of effective schools devote too little attention to the role of students, parents, and citizens in establishing and maintaining good schools.

Need for Better Measures of Student and Teacher Perceptions

We have argued that measurement of classroom processes--a careful examination of what happens in classrooms--has made the school effectiveness

literature more valuable. We also believe that information about how students and teachers perceive instructional processes and opportunities in more effective schools is needed to provide clues about how to make schools more effective. Although there have been fruitful attempts to measure general reactions of teachers and students to school (e.g., Brookover et al., 1979), future research needs to focus more closely on participants' reactions to specific events, especially events believed to be central to school effectiveness.

It is also important to assess the influence of school culture and instructional processes on students' perceptions. It is one thing to say that students and teachers should hold high expectations, but another to get answers to specific questions: How do students know how hard they should work? How can students know whether they are devoting more or less effort to schoolwork than their peers? The recent work of Natriello and Dornbusch (1984) illustrates the value of measuring student perceptions as well as the difficulty of doing the work.

Measurement of students' perceptions and observations of what they do in classes should be more central to the study of effective schools than it has been in the past. There is growing evidence that student perceptions of classroom process are valuable sources of information about schools (e.g., Cooper & Good, 1983; Peterson & Swing, 1982; Rohrkemper, 1984; Weinstein, 1983; Wittrock, in press); however, student data are more useful when combined with process observation and contextual information. Future studies of effective schooling could make better use of student interviews in order to understand how different types of students perceive and act upon the various constraints present in more and less effective schools.

Teacher Perceptions

Similar arguments could be made about the value of measuring teacher beliefs, perceptions, and decision-making skills related to effective schooling (see Clark & Peterson in Wittrock, in press). Why do some teachers believe that students can achieve curricula goals while other teachers in the same school do not? Why is it that more teachers in some schools believe that students can learn than do teachers in other schools serving similar populations? There are important data to suggest that teachers' expectations for student performance vary from school to school (see Brookover et al., 1979); however, needed now are assessments of other teacher perceptions that may help to explain why some teachers hold high expectations for student learning. Cusick (1983) provides a good description of schools from the teacher's viewpoint.

School Effects

Despite the considerable attention currently paid to school effects, few researchers have conceptualized and operationally defined school effects. As noted above, in defining an effective school, researchers commonly use only a single grade level, and in studying school process they observe only a small portion of teachers in a school. Thus, observation (when it does occur) usually focuses on the classroom rather than the school. Although some studies have examined school-wide practices (e.g., school rules), many important school measures have not been examined. For example, are teachers in effective schools more aware of what other teachers do in their classrooms? Do teachers in these schools have more opportunities to learn from other teachers (e.g., to observe, to engage in formal discussion) or to receive useful feedback from them? If teachers receive more feedback, what is the nature of the feedback? Is there a shared language and emphasis among staff in how they

assess the school's present strengths and weaknesses and how teachers perceive and define problems that confront the school (e.g., Spencer-Hall & Hall, 1982)?

Independence of School Effects

Researchers will have to conceptualize and isolate school effects more carefully than they have in the past. Effects on individual students are averaged to produce mean teacher effects; effects of individual teachers are averaged to produce mean school effects; and effects of individual schools are averaged to produce mean school district effects or estimates of the general effects of schooling. At each level of aggregation the variance within units probably exceeds the variance across units, so that prospects for meaningful findings at the higher levels of aggregation are limited. Attention to individual process variables makes it possible to logically separate effects on students that result from individualized (dyadic) teacher behaviors, teacher behaviors directed to the entire class, school effects representing actions of the principal or other sources of influence on the school as a whole, school district-level effects, and so forth. However, the effects of all of these on achievement are confounded in reality.

District Effects on Schools

There are a variety of ways in which school districts may encourage or restrict school effectiveness. For example, in some school districts the district office controls most money and uses it, for example, to employ curriculum coordinators throughout the district. Although individual schools may vary in the extent to which they take advantage of coordinators' skills, the variation is probably not as great as it would be if the money were available for schools to pay their own coordinators. For example, if principals were given funds to employ instructional leaders in their schools, the effects

on schools of using these leaders would probably be more apparent. Some schools would use those resources in practical ways (observing teachers and providing meaningful feedback, developing effective programs for identifying and training substitute teachers--whereas other schools might use the funds in less helpful ways (clerical tabulations).

Thus, the potential for school effects is larger in some districts than others. This observation is important, because some districts are currently decentralizing policy (e.g., individual schools set their own goals) and some school boards are redirecting control of resources from the district to the school level. As we noted above, influences at every level are confounded by resources and practices at other levels; hence, in districts in which individual schools have more power (e.g., teachers have some influence on whether principals are retained, principals can actively recruit teachers), and in which individual schools are allowed to determine how funds are spent, differential school effects are more likely to emerge.

Despite the potential impact of the district on school policy, principal leadership, and classroom processes, there has been little formal study of these effects in the past. There is now growing empirical work (see, for example, Bidwell & Kasarda, 1980; Schwille et al., 1983) and logical arguments (see Dreeban & Barr, 1983); however more work is needed in this area (in particular it will be important to show how decisions reached at the district level affect resources and instructional treatments that individual students or classes receive).

It is also possible to explore how the larger effects of the community or state on decisions and actions in particular schools. Hall and Spencer-Hall (1980) describe a case study that illustrates how an article that appeared in a newspaper was instrumental in forcing a school district to examine unanticipated issues in the area of testing policy and accountability measures. More

broadly, Spencer-Hall and Hall (1982) provide a detailed analysis of two school districts they observed for about a year. They found that school problems identified in the two school districts varied in their source, duration, resolution, and negotiation. Differences between the districts in perceiving and responding to school problems could be explained in part by community context, organizational history, organizational complexity, participant efficacy, and teacher militancy. Clearly, individual schools are affected by the larger organizational and political context that surrounds them.

Parental-School Relationships

The degree of home and school cooperation is likely to be an important determinant of student achievement. However this "obvious" possibility has received little research attention. Whether parent-school communication differs in "more" and "less" effective schools is also unclear. One might predict that home-school agreement on curriculum content and performance standards would influence student performance. However, there is also reason to suspect that parent involvement in and understanding of school programs is a complex matter.

The difficulty of understanding the potential value of parent-school relationships is due, in part, to the fact that this relationship can take many forms: understanding and supporting the school curricula programs or discipline policies; helping students with homework; raising money for the school; participating directly in the instructional program. As the definition changes, both the perceived (by parents and teachers) and actual benefits of parent involvement may vary considerably. Unfortunately, extant research does not provide a clear analysis of the forms nor effects of these different parent involvement activities in "more" or "less" effective schools.

Lightfoot (1978) argues persuasively that productive home-school relationships require that parents and teachers to see and accept the fundamental home and school to base their interaction and planning on goals of integration and cohesion rather than relationships that emphasize boundaries and individual interests. Whether the views of parents and teachers are more aligned on certain aspects of schooling in "more" than "less" effective schools is an intriguing issue that marits study.

The Role of the Principal in Program Adoption

To this point, we have reviewed research on effective schools and have made suggestions about subsequent research that needs to be conducted. We now turn to a discussion of the principal. Although the topic is logically interrelated with the issue of effective schools, we discuss the research on principals separately because research on principals has generally been conducted in a separate research tradition.

There is growing evidence that principals and other school-level factors may interact with the extent to which individual teachers adopt certain teaching practices to determine the success of program innovations. In this section of the paper we present two examples of research that show how principal behavior can affect program adoption.

A Secondary School Example: Stallings and Mohlman (1981)

Stallings and Mohlman examined the effects of school policies on pupil outcomes in an inservice training program in eight secondary schools that included a wide variety of policies and organizational plans.

They gathered opinions about school structure from school superintendents, county coordinators of secondary schools, and Stanford University consultants who had studied a large number of Bay Area schools. School structure variables included rule clarity, rule enforcement, communication patterns, leadership style, grading systems, delinquency, and a general feeling that a

a school was well run or not. After potential participating schools were identified, the investigators met with principals and arranged for a meeting with teachers at the schools of eight principals who volunteered to participate early in the school year. These principals were willing to commit staff development funds for release time if their teaching staff agreed to participate.

The initial school sample was balanced in terms of student ethnicity and average family income. Two of the schools had primarily white populations with students whose families had average or above-average incomes. The other six schools had large minority populations and family incomes that ranged from average to below average.

During the fall the investigators met with 10 to 12 volunteer teachers from each of the eight target schools. Three schools declined participation and three other schools replaced them. Four to seven teachers from each school ultimately participated in the training by attending seven one- to two-hour workshops.

To investigate relationships between school structure and organization, and between classroom teaching processes and school, teacher, and student outcomes, the researchers used a pretest-posttest design. Workshop training was designed to inform teachers about recent studies of effective teaching and to encourage them to use certain practices in their classrooms. Teachers were observed with a coding instrument that measured the extent to which they implemented practices believed to be correlated with student achievement. Other instruments used in the project included pupil and teacher questionnaires. The study also included interviews of principals that enabled the investigators to obtain more information about the school.

The investigators report the following major findings:

1. Schools in which policies and rules were clearer and more consistently enforced had higher teacher morale, fewer classroom intrusions, a lower absence rate, less class misbehavior, and more time-on-task.

2. Schools that had more administrative support services and fewer burdensome duties for teachers had higher teacher morale and less classroom misbehavior.

3. A more active and respected principal was associated with higher teacher morale and students who felt more friendliness.

4. In schools with more supportive principals, more teachers implemented the training program.

5. In schools where the policies and rules were clear and consistent, more teachers changed their classroom behavior as recommended.

6. Schools in which teachers implemented the effective use of time training programs had students who spent more time-on-task.

7. Findings regarding effective school policy and principal leadership style were similar for schools serving high-income and low-income students.

It is important to note that Stallings and Mohlman worked with volunteer principals and teachers and that these teachers were only a small percentage of the staff at each school. These data therefore do not yield strong statements about effective schools. Findings nevertheless provide good evidence that a principal's support is important in helping teachers to learn and use new instructional approaches, although some teachers in schools with less active principals also implemented the training information effectively.

An Elementary School Example: Hall et al. (1983)

Gene Hall and his colleagues at the Research and Development Center for Teacher Education have focused their research efforts on school organization and school change. Most recently they investigated the influence of principals as change facilitators in schools (for a review of this research see

Hall, Hord, Huling, Rutherford, & Stiegelbauer, 1983). One of their studies, the Principal-Teacher Interaction Study, was an extensive examination of nine elementary school principals and their faculties who were involved in implementation and innovation at their schools.

Although the complexity and scope of Hall et al.'s (1983) work prevent us from reporting it in detail here, it is important to note that they found that the principal played a key role in successful implementation of change strategies and that successful innovation called for much time and energy from the principal. However, they also found that various principals used different leadership styles successfully. They distinguished three principal styles: *initiators* (make it happen), *managers* (help it happen), and *responders* (let it happen).

Such data suggest that there is *no* ideal style of principal leadership, even in the specific role as a change facilitator. Principals who use any one of these three general styles can be successful. Needed now are data on the quality of leadership style (How do poor initiators differ from good ones?) and the contexts in which a given style may be more or less appropriate.

An Experimental Study

There is reason to believe that principals' involvement in change efforts can make a difference in program adaptation at the elementary school level. Gall, et al. (1984) recently completed a field experiment to determine if the effectiveness of a staff development program for elementary teachers could be improved by training the teachers' principals in a staff development program. Although instructional leadership is considered a key skill for principals to develop, attempts to assess principals' impact on instructional programs have been rare. This project provided evidence that trained principals, on average, had a positive impact on teachers' implementation of the Missouri Mathematics Program. This study provides the first *experimental* evidence the

principals' leadership behavior affects teachers behavior and student achievement. For an excellent discussion of the principal's role in staff development as well as a more general discussion of the role of staff development in the change process, see Gall et al. (1984).

The Principal as Leader

Although much school effectiveness research suggests that principals' involvement as curricula and instructional leaders is crucial, few existing data *describe* what principals do or tell how principals in more effective schools differ in their behavior from principals who head less effective schools. Fortunately, research has begun to focus on principals' behavior, and in the next decade there will likely be many observational studies of principals.

Case Studies of Five Principals: Dwyer et al. (1982)

Dwyer, Lee, Rowan, and Bossert (1982) argue that to understand leadership it is necessary to examine school processes because correlational survey studies cannot provide information about *why* some schools are more effective than others.

Dwyer and colleagues studied five principals for eight weeks. These principals were selected from a larger group who had been identified as effective by superintendents and other central office administrators from Bay Area school districts. Thirty-two principals agreed to be interviewed extensively. Five of these who were most articulate in describing their jobs and seemed most interested in the study were selected for participation in the case studies. To be certain that principals of effective schools were studied, the investigators prepared seven-year school achievement score profiles. They believed that such data at least allowed them to exclude ineffective schools (in terms of achievement) from the sample.

Open-ended interviews were conducted in order to determine principals individual philosophies of work observed in an unobtrusive manner, and decriptive field notes were kept of their activities. The five principals were observed during three full work days over an eight-week period, with these days sometimes consisted of 10 to 12 hours of observation. The day following each observation, the researchers returned and interviewed the principals about the previous day. Principals were encouraged to reflect on their decisions and activities. In addition to the interviews and observations, researchers spent 20 to 30 hours in each school observing classes, recess, lunch periods, and talking informally with teachers and students. Information such as school plans, test scores, and so forth were also obtained from each school.

Data analysis was similar to the "comparative method" framework of Glaser and Strauss (1967), with analysts generating definitions and categories from the records, searching for patterns, repetitions, and contradictions in each setting, and comparing the obtained results across settings. Narrative case studies were prepared and summarized in models that illustrated the essential qualities of a school's context, the activities that best typified the principal's management behaviors, and the expected outcome of those actions as projected by the principal. These models were then discussed with the principals in order to assess accuracy and were modified or verified on the basis of these discussions.

All the principals believed that their personal backgrounds influenced their school activities, and the observations mostly bore out these claims. For example, the democratic and egalitarian beliefs of one principal were consistent with the way in which he worked with school faculty. Dwyer et al. believe that such principal characteristics have implications for training because such knowledge could be used to match principals to schools that need specific types of leaders.

Principals' reports and the observations indicated that the community also had a significant effect on the behavior of principals. For example, one principal noted that he spent 60% of his time responding to situations that orginated in the community. On the one hand, principals viewed their communities as constraining influences, or something that took time from tasks they would rather perform. However, principals also saw that the community could provide both materials and personnel, areas in which schools face serious shortages.

This study implies that institutional context both limits and provides opportunity for principals, and various administrators may react very differently to the same programs, pressures, and opportunities. In this study, principals' reactions to district-level programs varied most. Some principals were skeptical about including district goals and efforts in their school plans; other principals considered this procedure useful. Dwyer et al. found that the context of a large school (e.g., student turnover, funding cuts) often complicates the work of a school principal. However, this effect is mediated by the principal's characteristics.

In addition to principals' personality and training, situational factors also affected principals' behavior. For example, the two principals who were least obtrustive in instructional matters led faculties who had taught for 10 years or more. The more direct and intervening principals led less mature teacher faculties or ones in which more turnover occurred. Successfully leading stable, experienced teachers, then, may require different strategies than leading inexperienced teachers.

The observers reported that all five principals were active and that they were worn out after following the principals around during the day. (One

wonders, however, if these principals are similar to most others.) The researchers state:

> These five principals also seemed to respond to daily cycles within their schools. First, they roam their buildings as children arrive, assessing potential problems and making sure classrooms are staffed and ready for the day. Next, they return to their offices for short-term planning, telephoning community leaders, and receiving the first round of student problems, which at this time of day are frequently related to situations in the students' homes. Then they move once again, to tour the building as recess begins, monitoring, solving problems, communicating with staff and students as they patrol. Between recesses and lunches, they commonly remain at large in the building observing classes, again talking with students and teachers as they move in and out of classrooms. Lunch periods and the hours following are frequently consumed by disciplinary problems which require interaction with students, teachers, or contacting parents with bad news. Dismissal at the end of the student day again brings these principals back to the hallways and public spaces of the building, where they admonish or praise, prompt or prohibit in rapid-fire encounters. The ensuing relative calm allows time for reflection and follow-up parent conferences, teacher conferences and staff or committee meetings of all sorts. (Dwyer et al., 1982, pp. 53-54)

These investigators suggest that daily activities enable the principal to assess how the school is functioning, to react to student misbehavior appropriately, and to suggest changes in teacher style or demonstrate new teaching styles to teachers. They speculate that the effects of these routine behaviors on the quality of instruction and on students' experiences can be substantial. In general, they found that all five principals had theories that guided their actions. All sought to understand how modifications in the structures of their schools influenced students, and all believed that their activities could and did affect how students learned.

Dwyer et al. contend that the study of these five principals produced a view of instructional leadership as accruing from routine, mundane acts related to the principals' perspectives on schooling. Furthermore, on the basis of their analyses, they argue that there are no simple ways to understand the

effects of principal behavior on schools and that more process studies are needed of principals' behavior.

This study and Wolcott's (1973) examination of one principal are the most intensive, comprehensive efforts made so far to describe in detail what principals do in their day-to-day activities. Because of their small sample sizes, however, these studies are more successful in raising issues than in resolving them. For example, the principals that Dwyer et al. studied had coherent visions of what a school should be, and they attempted to achieve these goals in their daily work. However, sampling characteristics probably made it likely that relatively more aggressive, active individuals were selected for the case studies (the five effective principals were selected largely because of their interest in the study). The energy the principals expanded was impressive--they were *very active*. However, there are no data documenting that ineffective principals are less energetic and more reactive than proactive (it is too bad that average or relatively ineffective principals were not studied in this research).

Nevertheless, the Dwyer et al. study is an important initial examination of the principal's role through *observation*. At a minimum, their data suggest that there are different ways in which principals may be effective and that future research needs to examine simultaneously principals' personal beliefs, characteristics of the schools they serve (type of pupils, maturity of teaching staff), community and school district variables, and most importantly, how principals mediate conflicting sources of pressure. The work by Dwyer et al. is still in progress and as further analyses are completed and their findings are more integrated it may be possible to devise a more concrete, specific picture of the role of a relatively successful principal.

Principals in High- and Low-achieving Schools

Brookover et al. (1979) conducted a less intensive study of four principals in two high-achieving and two low-achieving elementary schools. These researchers' major purpose was not to examine the principal's role in effective schools, but rather to determine if and how schools made a difference in student achievement (see earlier discussion of their work in this paper).

Four case studies were completed in order to compare pairs of schools that had similar racial composition (two predominantly white; two predominantly black) and SES levels and were situated in comparable communities but had different achievement levels (in each pair one school was comparatively high in achievement; the other school had a poor performance record). Achievement was measured by a school's mean score on the Michigan Assessment Test for 1974. Because schools within each pair were matched closely on demographic variables, it seems plausible to attribute differences in school achievement to social and process variables within schools.

Unfortunately, Brookover et al. do not describe the observational variables or procedures they used, so it is difficult to analyze their narrative reports. More information about the type of training observers received, the number and duration of observations, and so on would allow broader conclusions about the stability of the data reported. With this qualification in mind, the major findings concerning the principal's role in each school are discussed below.

The principal in the high-achieving, largely white school indicated that his primary concern was student achievement. He organized his work so that an assistant completed most of the administrative paperwork, leaving the principal time for instructional supervision. The principal's emphasis on

instruction was perhaps best demonstrated by the announcement at a staff meeting that he intended to observe every teacher at least 30 times (as he had done the year before). He encouraged teachers to participate actively in in-service meetings and to discuss with him ways to improve the school. Furthermore, he wanted them to ask him for instructional advice. The researchers report that some of the teachers, especially the older ones, preferred that the principal be primarily a disciplinarian and spend less time on classroom activities. Still, all teachers seemed to respect the principal and to recognize and appreciate his interest in student achievement.

The principal in the high-achieving, largely black school was primarily an effective administrator who kept good records. He supported teachers and encouraged them to improve their instruction by attending inservice programs. Teachers in the school, however, attributed the school's effectiveness to a principal at the school four years earlier who had been an instructional as well as an administrative leader. For example, the former principal often led inservice training instead of merely encouraging teachers to attend. Still, the current principal periodically observed and critiqued teachers' classroom instruction; however, he believed that the primary responsibility for the quality of education in the school rested with the teachers.

The principal in the low-achieving, largely white school was the administrator of two schools and hence was only in this school on alternate mornings and afternoons. He viewed his major responsibility as dealing with problem students, and there appeared to be little interaction between the principal and the teaching staff. He spent much of his time compiling files on students with behavior problems and working with these students and their parents to solve these problems. The principal held low performance expectations for students. Furthermore, he rarely observed in classrooms, spent much time in

the public areas of the school, or visited with teachers or students. In short, this administrator offered virtually no academic leadership.

The principal in the low-achieving, primarily black school served mainly as an administrator and disciplinarian. However, unlike the principal in the low-achieving, primarily white school, she was not aloof from teachers; indeed, she interacted frequently with the staff and spent much time in the teachers' lounge. At the same time, this principal had low expectations for most students. Although she expressed interest in students' achievement, she spent little time observing or critiquing teachers, and did not seem to expect a high level of teacher performance. Perhaps the most telling evidence of her lack of instructional leadership was the fact that the assistant principal was primarily responsible for observing teachers and keeping academic records, but had no power to effect change.

Results of the Dwyer et al. and Brookover et al. studies indicate that some relationship exists between the principal's role expectations and student performance. Perhaps the best way to express this relationship is that appropriate expectations on the part of a principal that students can and will learn the curriculum are necessary, but not sufficient, for effective teaching and learning to occur in a school. In addition, principals must be aware of and concerned about classroom instruction. Although it may not be necessary that principals assume direct leadership of the instructional program (e.g., Gersten, Carnine, & Green, 1982), at a minimum, they must make certain that teachers obtain information about their classroom teaching and feedbsck about the adequacy of their instruction.

Policy Implications of Principal Effectiveness Research

Although we have focused on two studies because of their attention to process measures, other recent studies also examine principal behavior.

Indeed, Manassee (1985) reviews research on principal effectiveness and comments on the policy implications of such work. She notes that recent research (e.g., Martin & Willower, 1981; Morris & Crowson, 1981; Pitner, 1982) provides a useful description of the day-to-day behavior of principals. It is clear from her review that principals' administrative work often includes tasks that are initiated by others, of short duration, frequently interrupted, and done face-to-face interaction (e.g., little written communication). Considering this environment, it is little wonder that principals develop a preference for dealing with concrete, immediate, and potentially solvable problems (e.g., discussing the agenda for tonight's PTA meeting over more distant goals and more problematic activities (e.g., trying to understand the quality of teacher or student interactions). Manassee argues convincingly that the policy imperative of recent research is the need to prepare principals for their fragmented, varied, and ambiguous role and to help them develop analytical, communicative, and instructional skills in order to provide more effective leadership.

Future Research on the Principalship

Although recent studies of principals and their effects on teachers and students are encouraging (recent studies have started to examine what principals do), researchers have not examined many important aspects of the principal's role.

Context

It may be that a different type of principal is needed to *improve* a school's achievement than to *maintain* an already adequate school achievement record. More information is needed about the principal's role in different

types of schools and about the duties of the principal in reaching goals other than those associated with basic academic achievement.

Assignment of Students and Teachers

How principals assign students to classrooms likely has a significant effect on what can be accomplished in a classroom or at a grade level within a school. Unfortunately, this variable and related student composition issues (e.g., How many honors sections will there be in a particular subject?) have been poorly defined and seldom studied systematically.

There is empirical evidence that classroom composition (the mean and range of students present in a class) affects student achievement. For example, Beckerman and Good (1981) studied the ratio of high- and low-achieving students in classrooms from a large metropolitan school district that served a middle-class population. They defined classrooms with *more favorable* teaching situations as those in which more than a third of the students were high aptitude and less than a third were low aptitude. *Less favorable* classrooms were those in which less than a third of the students had high aptitude and more than a third were low aptitude.

Beckerman and Good found that both low- and high-aptitude students in favorable classrooms had higher achievement scores than the two groups in unfavorable classrooms. Veldman and Sanford (1982) also found evidence that classroom composition is associated with student achievement. They measured classroom composition in nine junior high schools by determining the mean achievement level for each class at the beginning of the year. They found significant interaction effects, indicating that both high- and low-ability pupils do better in high-ability classes and that the effects of class ability are more pronounced with low-ability students.

Others, too, have recently explored the context effects of mean classroom ability and degree of heterogeneity on classroom outcomes. For example, Evertson, Sanford, and Emmer (1981) found that heterogeneity in students' entering achievement levels restricts classroom teachers' ability to adapt instruction to individual students' academic and affective needs. Higher heterogeneity was also associated with a lesser degree of student task engagement and cooperation. Leiter (1983) found, as did Beckerman and Good (1981), that students receiving instruction with high-ability classmates made substantially higher mathematics gains than did those with low-ability classmates. However, this effect did not hold in reading. Rowan and Miracle (1983) explored the effects of ability grouping on the achievement of fourth-grade students in a single urban school district. They found that both within-classroom grouping for reading instruction and across-classroom ability grouping had direct effects on reading achievement, and demonstrated that ability grouping tends to reinforce initial inequalities. (However, certain aspects of within-classroom grouping had favorable effects for students in the low group.)

Dreeban and Barr (1983) demonstrated that assignment of students to classes can affect how teachers organize classes for instruction. In a study of 15 first-grade classes, they found that in small classes (of about 20) with few low-aptitude students, teachers began with whole-class instruction and later reorganized classes into heterogeneous groups. Large classes (with 30 or more members) differed, both in their initial grouping arrangements and in subsequent modifications, according to the number of low-aptitude students they contained. Teachers who had fewer low-aptitude students employed a wider variety of group arrangements. Teachers with many low-aptitude students inevitably created a large low group as part of a classroom figuration

consisting of three groups of equal size. The mean aptitude level of each group was the major influence on the critical instructional decision of how much material to cover over a given period of time--the pace of instruction--which in turn was a major determinant of individual learning.

It is clear that placement in a class or group with less capable students makes it more difficult for a given student to make educational progress than placement with more capable students does. Yet grouping by ability or achievement level probably is necessary under certain circumstances. The assignment of students to classes is a very important issue, though this topic has received little research attention.

Principal's Role in Assignment of Students to Classes

Although extant research has not achieved any systematic understanding of which principal beliefs lead to the assignment of more and less favorable classes, anecdotal evidence indicates that principals' beliefs about classroom learning are associated with their actual assignment of pupils to classes.

For example, Good and Marshall (1984) note that because of declining student enrollment in some American schools, students are grouped across grade levels in order to have sufficient numbers of students for a class. It is likely that some principals' grouping of students is influenced more by organizational or institutional needs than by concern about how best to educate students.

One of the authors observed the effects of such decision making on the school lives of some students in a small school serving a diverse population. There were enough second- and third-grade students to justify the formation of three classes (one mixed, one second-grade, and one third-grade). In this

particular case, the principal decided to form the mixed class on the basis of student *maturity* (capacity to work independently) as opposed to *ability*.

The principal wanted mature third- and second-grade students in one classroom so that one group could work independently while the teacher worked with the other group. Had the principal formed classes according to ability, there probably would have been more pressure on the teacher to use whole-class and large-group teaching. Had the principal used more dynamic individual characteristics (sociability, works well in groups), or stressed a more social outcome (learn to work well with others who are diverse), the teacher might have made greater attempts to have second- and third-grade students interact.

In this case, the independent worker model and the demand characteristics communicated to the teacher by such a grouping virtually guaranteed that the teacher would instruct the second- and third-grade students as separate, *intact* groups (without much social or academic contact between groups), and that comparatively little social interaction would be allowed *within* groups because group work was institutionalized as individual work.

This class contained 16 second- and third-grade students. The four third-grade girls appeared to be socially isolated, in part because of peer expectations (i.e., social interaction occurs with same-sex, same-age classmates and the teacher did little to alter this peer norm), and in part because the girls were from diverse backgrounds.

This example clearly illustrates the need to study a variety of variables if classroom life is to be understood more fully. It is likely that the principal's decision about how to assign students was influenced to some extent by his perception of the teacher's style and ability, and that the teacher's classroom strategies were influenced by her assumptions about the principal's motivation in assigning this particular composition of students.

Another teacher or a different four girls might have led to different consequences.

The assignment of students to classes in secondary schools is also an important consideration. For example, in creating homerooms should a principal in a junior high or middle school with four to six feeder schools deliberately mix students from different elementary schools? What are the consequences of being an "outsider" assigned to a homeroom where most students are from the same elementary school (e.g., What are one's chances of being elected to office)? If all band members miss first-period classes, or if foreign language is taught only in the second and third periods, what consequences does this have for the composition of other classes (i.e., Does it make some easier to teach than others?). If a high school principal decides to have an honors section of mathematics, what effects does this have on instruction in other math classes? If students are assigned to sections on the basis of ability in high schools, which teachers should teach high and low sections?

Considering that well-qualified teachers in some areas (e.g., math and science) are scarce at many schools, how should these teachers be utilized? Should they teach half of the time and spend the other half working with other teachers? Should they teach beginning or advanced sections? Atkin (1983) notes that researchers have neglected the question of how truly talented and well-trained teachers should be assigned and deployed. Unfortunately, there is little information about strategies principals use to assign teachers and students to classes or about the consequences of those strategies.

Feedback to Individual Teachers

Researchers should also examine how principals influence instructional behavior in their schools. How do they communicate expectations and establish

instructional priorities? If principals encourage teachers to determine their own instructional goals, how do they become aware of each teacher's goals, and how do they monitor and provide feedback about progress? Dwyer et al. (1982) provide a helpful profile that characterizes effective principals as being *more active*. Considering that effective principals are more visible to teachers and students in their schools, it is important to know if the *quality* of their decisions and actions can be related to student progress. For example, when and how often do principals visit particular classrooms? Do successful principals spend more time with teachers they believe to be average teachers, or do they observe less capable teachers more closely? How specific is the feedback they provide to teachers? On what topics do conferences focus? Do more effective principals discuss issues of curriculum and instruction, or do they talk only about general issues of classroom management, resources, and human relations?

Because of our past interest in teacher expectation research, we have a special curiosity about principals' communication of low expectations to certain teachers. Just as some teachers expect too little from certain pupils, some principals likely expect too little from certain teachers. It would be valuable to study how principals communicate expectations to teachers (e.g., the classes or students assigned to a teacher, the committees or duties assigned, the way requests for supplies are handled, the frequency and degree of formality of observational visits, etc.) and to determine how principals vary in their ability to communicate in positive and helpful ways with teachers.

In the 1970s and early 1980s, useful literature for describing teacher planning and thinking has developed (see Clark & Peterson, Chapter 11 in Wittrock, in press; Shavelson, 1983). Hopefully, the 1980s will see the development of more systematic knowledge about principals' decision making and

behavior. We have discussed but two of the many areas that future research could contribute to. Needed are more detailed studies of what principals do, of why principals act as they do, and the apparent effects of their beliefs and behaviors. Research thus far indicates that effective principals are more proactive (although many aspects of proactive behavior have not been examined e.g., Do more successful principals watch teachers teach before hiring them?) and visible, although more research is needed to validate this viewpoint.

A Synthesis of Research on Effective Schools

Nearly all studies of effective schools support the importance of principal leadership. There is far less consensus, however, on the behaviors and practices that characterize leadership on a day-to-day basis. Ironically, while principals tend to rank instructional leadership as their most important function, available evidence suggests that they have little time or opportunity to provide such leadership (Howell, 1981; McLeary & Thompson, 1979).

Rather, as Cohen (1983) notes, the work of principals is characterized by ambiguous and conflicting expectations, frequent interruptions, and crises. Principals tend to engage in short tasks or brief interactions, often as many as several hundred per day. Their interactions tend to be personal and problem centered (Morris, Crowson, Hurwitz, & Porter-Gehrie, 1981; Wolcott, 1973). Further, Salley, McPherson, and Baehr (1979) describe principals as captives of their environments, strongly influenced by the structure and organization of the school and the school district.

How, then, do principals affect school achievement? First, their goal orientation is especially important (Bossert, Dwyer, Rowan, & Lee, 1982; Greenfield, 1982). Cohen (1983) argues that effective principals emphasize

achievement, set instructional goals, develop performance standards for students, and express optimism about the ability of students to meet instructional goals. He notes that, not surprisingly, in light of the inherent constraints of the role, effective principals need to be proactive, to develop and articulate a vision of the school and its future, and to project that vision in the course of numerous daily interactions with teachers (Blumberg & Greenfield, 1980; Little, 1981a, 1981b, 1982).

Compared with less effective principals, effective principals tend to take responsibility for instruction, observe teachers regularly, and discuss their work problems. It is clear that some principals have little awareness of what takes place in the classroom, so that there often are large discrepancies between recommended district policy and actual classsroom practice (see, for example, Ebmeier & Ziomek, 1983). Principals can promote effective teaching by creating the conditions that enable it to occur, and by preventing or limiting intrusions once it is underway.

Cohen suggests that although it is possible to describe some of the things that principals do to contribute to instructional effectiveness, it is clear that research does *not* tell us that all effective principals engage in all of these activities, nor does it yet tell us about the conditions under which certain strategies are likely to be more or less appropriate or effective. To date, only a vague outline of a rather complex picture has emerged, and much more research needs to be done to complete the picture (see also Little, 1982). Principals must do more than provide instructional leadership; their ability to create shared *values* and *culture* is also important. Presumably, principals in effective schools generate a strong sense of community, with commonly shared and high expectations for student and staff performance. Cohen notes that community in a school requires more than shared

instrumental goals; it requires the creation of a moral order that entails respect for authority, genuine and pervasive caring about individuals, respect for their feelings and attitudes, mutual trust, and the consistent enforcement of norms that define and delimit acceptable behavior (see also Grant, 1982). The importance of a shared moral order should not be underestimated, because schools are fragile social institutions easily disrupted by conflict in or around them. For example, there are weak formal controls over the selection of staff, and students are involuntary clientele of the school.

Cohen argues that student and faculty norms and school "ethos" can be shaped by principals and teachers, as well as by several structural features of schools. One feature, building-level autonomy, refers to the view that circumstances among schools, even within a single district, vary considerably. Schools differ in their mix of students and staff; the characteristics of the communities they serve; the histories of their attempts at innovation and improvements; the prevailing norms, beliefs, and shared understandings; and the problems they face. From this point of view, attempts at instructional improvement will be successful only to the extent that schools are given sufficient latitude to adapt new policies or practices to their unique circumstances or to develop their own solutions to problems.

A second structural feature involves procedures for assigning students to schools. Cohen notes that the advantage some private schools appear to have in terms of creating appropriate climates may result from their procedures for recruiting new students. However, as we argued earlier in the chapter, how a principal assigns students to individual classes influences what can and does take place in the classroom.

A third structural feature of schools suggested by findings of research on school culture relates to the quantity and organization of time in schools.

Simply put, shared work and collective decision making require time for teachers to talk with each other, to observe each other's classrooms, and to plan and evaluate programs. Cohen points out that policies that lengthen the school day or year could be used to free groups of teachers from classroom responsibilities during part of the school day in order to create additional time for planning and shared work.

In his conclusion, Cohen states that effective schools become so by making headway in solving several problems that are rooted in the structure of educational organizations and the teaching profession. First, because students are involuntary clientele who bring conflicting goals to school, there needs to be a relationship of warmth and trust between teachers and students. Teachers in effective schools apparently are able to bring about the feeling in students that school achievement norms are legitimate.

Second, by their very structure, schools serve multiple social functions that compete with their instructional mission. Additionally, over the past several decades, the range of social concerns that schools have been asked to address has increased considerably. Consequently, schools are seen as having diffuse goals and little cohesion (Good, Biddle, & Brophy, 1975). Effective schools, however, have been able to assert the primacy of their instructional mission around a limited set of goals, and in ways that direct and focus the allocation of resources, operating procedures and practices, and the behavior of teachers and students. The means for improving teaching effectiveness have therefore been quite limited, for without the knowledge to relate means to ends, choosing the most productive among alternative practices is difficult.

Cohen believes that the quality of research conducted over the past decade has changed this condition considerably. Practices for motivating, instructing, and controlling students in classrooms are better understood.

There is now a sizable knowledge base now describing teaching practices that can be related to student learning (see Brophy & Good, Chapter 10 and Rosenshine & Stevens, Chapter 13 in Wittrock, in press). Although there is still much that is not understood about teaching, the important point is that what is known is also used in effective schools, so that the technology becomes more explicit and precise, and choices about alternative practices are made more knowledgeable.

Finally, teaching is a profession in which work is typically performed in isolation from one's colleagues. Cohen suggests that this isolation has several undesirable consequences, including the limited codification of successful practices and a tendency for teachers to treat uncertainties inherent in the role as personal, rather than collective, problems. By developing collegial working relationships, effective schools tend to alleviate this problem and its consequences, and teaching becomes shared work. Teachers can thus learn from one another as well as distinguish limitations inherent in the profession itself from those related to an individual teacher's capabilities. The result is both improved teaching practices and enhanced professional self-esteem.

Summary

Prior to the mid 1960s there had been relatively little observational research that examined schooling in America. The publication of the Coleman et al. report (1966) initiated a flurry of input-output studies that attempted to examine whether school resources (e.g., the ratio of adults to children; the number of books in the library) were associated with student outcomes (typically performance on standardized achievement tests). These studies generally did not show any consistent relationship between resources and student outcomes (Averch et al., 1974).

This paper has focused on the generation of studies that followed the input-output efforts and examined studies that include measures of school process, school input (quality of student body or teacher staff; school resources) and school outcomes. The latter research yields findings different from those of input-output studies. First, it shows that processes associated with individual schools are related to student achievement, and suggests mechanisms through which some schools obtain more achievement than do other schools that have similar inputs. A second major finding is that some processes consistently characterize more and less successful schools. Although studies can be faulted because of methodological problems, the fact that several studies reach similar conclusions lends credibility to the claim that certain processes are associated with school effects on achievement.

Despite this consensus, however, school effects data are limited in several respects. First, most effective schools research has been conducted in urban schools, so its application to suburban schools is unknown. Second, the description of effective schools is based largely on the schools' effectiveness in obtaining high student performance on standardized achievement tests. This is a narrow definition of school effectiveness. Although there is some evidence that schools can simultaneously achieve several goals (e.g., high attendance rates, high student engagement rates, high achievement), for the most part the question of school success on cognitive criteria other than standardized achievement (e.g., decision-making skills) has been ignored. There is no evidence that schools that teach the basic skills relatively well can also teach computer skills, science, and writing relatively well. Furthermore, process measures usually have been limited to a few global dimensions of schooling, and these examine *form* more

than *quality*. Often, data are collected on only a few teachers per school, and the information about what even these teachers do is sketchy.

Future research on effective schools could better conceptualize and study school-level processes. For example, certain school practices and school-wide beliefs may make it more likely that teachers will order and use scientific equipment or take students on field trips. Rowe (1983) argues that insurance, permission, and supervision issues associated with field trips and red tape and safety considerations necessary when ordering, maintaining, and using laboratory equipment restrict teachers' use of field trips and laboratory work. It is also possible that in some schools certain administrative practices or beliefs about normative practice (the importance and necessity of lab work) may affect the frequency with which teachers take field trips and use laboratory equipment. If so, how and why are these administrative practices and beliefs created, and what maintains them?

Researchers might also ask whether in some schools it is easier for teachers to receive relevant feedback about their classroom performance from the principal or peers. Why do teachers in some schools receive more frequent and meaningful feedback about their instructional program, its consequences, and possible alternative courses of action than teachers in other schools? If in some schools teachers use other teachers more frequently as resources, what are the normative belief structures and administrative practices of those schools, how did they develop, and how are they maintained?

There is some research relevant to the issue of teachers receiving meaningful feedback about their classroom teaching (see e.g., Little, 1981b), and limited information about normative belief structures that operate in certain schools. Our point here, however, is that most effective schools research has not examined the school as a unit nor has process research been

organized in order to explore issues related to the school as a unit. Research in other areas has examined the school as a unit (e.g., Metz, 1978). What we need now is an integration of such thinking and conceptualization into the effective schools research.

Another major constraint on effective schools research is that existing evidence is largely correlational. Whether active leadership precedes or follows the development of high expectations or whether student achievement precedes or follows high expectations for performance is uncertain. Still, there is growing evidence that the effective schools literature can be *translated* into practices that improve student achievement. However, as we note in this chapter, this translation often goes far beyond what the literature indicates. As a case in point, most school improvement research advocates a written and formally agreed upon (by the school staff) plan for school improvement. Correlational research provides limited evidence that teachers and administrators have more of a sense of direction and a shared consensus about what is important in high-achieving schools than in schools where achievement gains are low. It is not entirely clear whether or not a formal school plan is equivalent to the consensus and direction reached by a teaching staff through informal dialogue and day-to-day communication built up over time. A school plan is not necessarily wrong or inappropriate, but under certain conditions, a plan may force too many accommodations and achieve only superficial compliance. Here we emphasize that correlational data must be translated into practice carefully. Extant data do not indicate that one should start with a school plan (perhaps the chance to observe other science teachers or to visit other schools, for example, should precede attempts to develop a related plan). One of our qualifications for school-wide change is that school plans should be constructed so that the initiative, autonomy, and

responsibility of individual teachers are preserved (the opportunity to make decisions about classroom events is one of the few things that makes teaching attractive for some teachers). Plans should also attend to the development of individual teachers (e.g., Good & Brophy, 1984; Spencer, 1984). Given extant enthusiasm in some places for master teacher plans and merit pay, and in other places for effective schools, it seems important to try to balance these two objectives.

Although it is not the focus of this paper, there is a growing literature on organizational theory and the change process in schools, and even some useful attempts to order and apply this literature for practitioners interested in improving schools (see Fullan, in press; Purkey & Smith, 1983b). Also, the school improvement efforts currently underway will provide clues about more and less desirable ways to translate extant research, as will the emerging literature that describes the behavior of school principals (Dwyer et al., 1982).

However, we need better *process* studies of school improvement implementation if we are to understand why some school improvement plans work when apparently similar plans fail. We also need more systematic attempts to develop theoretical explanations (see Biddle & Anderson, Chapter 8 in Wittrock, in press), and more basic research on other important outcomes of schooling. Although the experimentation now underway is useful, enthusiasm for application should not undermine attempts to investigate broader definitions of schooling. Research findings presently available *do not provide statements about effective schools*. The literature does yield statements about factors associated with raising students' performance on standardized achievement tests. Information about such effects is a useful step forward if educators

keep in mind that extant information about effective schools cannot be equated with effective schooling.

Ironically, in the past few months a number of national commissions have examined the state of American schooling. Many of their reports recommend mechanical solutions to complex educational problems. For example, the National Commission on Excellence in Education called for longer school days. However, one wonders how more time ipso facto can improve teaching and learning in American schools. Although reports like *A Nation at Risk* have been enormously successful in generating public interest--and potentially public support--for education, they have also shifted the debate on schooling from issues of quality (what teachers and students do in classrooms to allow for understanding and application of important knowledge) to the search for standard and immediate answers (more graduation requirements). As Bossert (1983), Hall (1983), and Slavin (1983) have argued (in response to *A Nation at Risk*), issues of *quality* must be addressed if we are to improve schooling.

In this paper we have argued that school officials need to be careful and to avoid overapplying school effects research. However, this research should not be ignored, as some policy advocates have done. It is clear that schools, even good schools, differ in subtle and important ways (Lightfoot, 1983). Calling for a longer school day in all schools because some schools need more time is a silly and costly educational prescription that will do as much harm as good.

Although certain aspects of the school effects literature may help practitioners to identify their problems and alternatives and thus allow them to think more systematically about their instructional programs, this research does not yield answers. The past decade has been an important start. The research completed to date shows that individual school variance is an

important dimension and that it can be influenced by selected actions and resources. Despite this progress, the next step does not involve application. Rather, it requires further extending the basic knowledge in this field by completing new studies that help us to understand more fully the qualitative aspects of schooling.

References

Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G. (1976). Analysis of the school preferred reading program in selected Los Angeles minority schools. Santa Monica, CA: Rand.

Atkin, J. M. (1983). The improvement of science teaching. Daedalus, 112(2), 167-187.

Austin, G. R. (1978). Process evaluation: A comprehensive study of outliers. Baltimore: Maryland Department of Education.

Averch, H. A., Carroll, S. J., Donaldson, T. S., Kiesling, H. J., Pincus, J. (1974). How effective is schooling? A critical review of research. Santa Monica, CA: Rand.

Beckerman, T. M., & Good, T. L. (1981). The classroom ratio of high- and low-aptitude students and its effect on achievement. American Educational Research Journal, 18(3), 317-327.

Berliner, D. (1977). Impediments to measuring teacher effectiveness. In G. Borich (Ed.), The appraisal of teaching. Reading, MA: Addison-Wesley.

Berliner, D. (1979). Tempus educare. In P. L. Peterson and H. J. Walberg (Eds.), Research on teaching: Concepts, findings and implications. Berkeley, CA: McCutchan.

Biddle, B. J. (1970). The institutional context. In W. Campbell (Ed.), Scholars in context: The effects of environments on learning. New York: John Wiley.

Bidwell, C. E., & Kasarda, J. D. (1980). Conceptualizing and measuring the effects of school and schooling. American Journal of Education, 88, 401-430.

Blumberg, A., & Greenfield, W. (1980). The effective principal: Perspectives on school leadership. Boston, MA: Allyn & Bacon.

Bossert, S. T. (1983). New nostalgia. Elementary School Journal, 84(2), 138-141.

Bossert, S. T., Dwyer, D. C., Rowan, B., & Lee, G. V. (1982). The instructional management role of the principal. Educational Administration Quarterly, 18, 34-64.

Brimer, A., Madaus, G. F., Chapman, B., Kellaghan, T., & Wood, R. (1978). Source of differences in school achievement. Slough, Bucks, England: NFER.

Brookover, W. B., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). School social systems and student achievement: Schools can make a difference. New York: Praeger.

Brookover, W. B., Erickson, E., & Joiner, M. (1967). Self-concept of ability and school achievement III. East Lansing: Michigan State University, College of Education (Cooperative Research Project 2831).

Brookover, W. B., Lapere, J., Hamachek, D., Thomas, S., & Erickson, E. (1965). Self-concept of ability and school achievement II. East Lansing: Michigan State University, College of Education (Cooperative Research Project 1636).

Brookover, W. B., & Lezotte, L. W. (1979). Changes in school characteristics coincident with changes in student achievement. East Lansing: Michigan State University, Institute for Research on Teaching. (ERIC Document Reproduction Service No. ED 181 005)

Brookover, W.B., Paterson, A., & Thomas, S. (1962). Self-concept of ability and school achievement. East Lansing: Michigan State University, College of Education (Cooperative Research Project 485).

Brookover, W. B., & Schneider, J. M. (1975). Academic environments and elementary school achievement. Journal of Research and Development in Education, 9, 82-91.

Califronia State Department of Education. (1977). School effectiveness study: The first year. Sacramento, CA: Office of Program Evaluation and Research.

California State Department of Education. (1980). Report of the special studies of selected ECE schools with increasing and decreasing reading scores. Sacramento, CA: Office of Program Evaluation and Research.

Clark, T. A., & McCarthy, D. P. (1983). School improvement in New York City: The evolution of a project. Educational Researcher, 12(4), 17-24.

Cohen, M. (1983). Instructional, management and social conditions in effective schools. In A. O. and L. D. Webb (Eds.), School finance and school improvement: Linkages in the 1980s.

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

Cooper, H. M., & Good, T. L. (1983). Pygmalion grows up. New York: Longman.

Cuban, L. (1983). Effective schools: A friendly but cautionary note. Phi Delta Kappan, 64(10), 695-696.

Cusick, P. (1983). The egalitarian ideal and the American high school. New York: Longman.

Doss, D., & Holley, F. (1982). A cause for national pause: Title I school-wide projects (ORE Publication No. 81.55). Austin, TX: Office of Research and Evaluation, Austin Independent School District.

BEST COPY AVAILABLE

Dreeban, R., & Barr, R. (1983). Educational policy and the working of schools. In L. Shulman and G. Sykes (Eds.), Handbook of teaching and policy. New York: Longman.

Dwyer, D. C., Lee, G. V., Rowan, B., & Bossert, S. T. (1982). The principal's role in instructional management: Five participant observation studies of principals in action. San Francisco, CA: Far West Laboratory for Educational Research and Development.

Ebmeier, H. H., & Ziomek, R. L. (1983, February) Student academic engagement rates (Final report) Washington, DC: National Institute of Education.

Edmonds, R. R. (1979). Effective schools for the urban poor. Educational Leadership, 37, 15-27.

Edmonds, R. R. (1982). Working paper. East Lansing: Michigan State University, Center for School Improvement.

Edmonds, R. R. (1983). Search for effective schools: The identification and analysis of city schools that are instructionally effective for poor children (Final report). East Lansing: Michigan State University, College of Education.

Eubanks, E. E., & Levine, D. U. (1983). A first look at effective schools projects in New York City and Milwaukee. Phi Delta Kappan, 64(10), 697-702.

Evertson, C. M., Sanford, J. P., & Emmer, E. T. (1981). Effects of class heterogeneity in junior high school. American Educational Research Journal, 18(2), 219-232.

Finn, C. (1983, June). Toward strategic independence: Policy considerations for enhancing school effectiveness (Final report). Washington, DC: National Institute of Education (NIE 400-79-0035).

Forsythe, R. A. (1973). Some empirical results related to the stability of performance indicators in Dyer's student change model of an educational system. Journal of Educational Measurement, 10, 7-12.

Freeman, D. J., Kuhs, T. M., Porter, A. C., Floden, R. E., Schmidt, W. H., & Schwille, J. R. (1983). Do textbooks and tests define a national curriculum in elementary school mathematics? Elementary School Journal, 83, 501-513.

Fullan, M. (1985). Change processes and strategies at the local level. Elementary School Journal, 85(3), 391-422.

Gall, M., Fielding, G., Schalock, D., Charters, W., & Wiczinski, J. (1984). Involving the principal in teachers' staff development: Effects on the quality of mathematics instruction in elementary schools. Eugene, OR: University of Oregon, Center for Educational Policy and Management.

BEST COPY AVAILABLE

Gersten, R., Carnine, D., & Green, S. (1982, March). Administrative and supervisory support functions for the implementation of effective educational programs for low income students. Paper presented at the annual meeting of the American Educational Research Association, New York.

Glaser, B., & Strauss, A. (1967). The discovery of grounded theory. Chicago, IL: Aldine.

Glenn, B. C. (1981). What works? An examination of effective schools for poor black children. Cambridge, MA: Harvard University, Center for Law and Education.

Good, T. L., Biddle, B. J., & Brophy, J. E. (1975). Teachers make a difference. New York: Holt, Rinehart & Winston.

Good, T. L., Biddle, B. J., & Brophy, J. E. (1983). Teaching effectiveness: Research findings and policy implications (Technical Report No. 319) Columbia, MO: University of Missouri, Center for Research in Social Behavior.

Good, T. L., & Brophy, J. E. (1984). Looking in classrooms (3rd ed.). New York: Harper & Row.

Good, T. L., Grouws, D. A., & Ebmeier, H. H. (1983). Active mathematics teaching. New York: Longman.

Good, T. L., & Marshall, S. (1984). Do students learn more in heterogeneous or homogeneous groups? In P. Peterson & L. Cherry-Wilkinson (Eds.), Student diversity in the organization process. New York: Academic Press.

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

Grant, G. (1982). Education, character and American schools: Are effective schools good enough? Syracuse, NY: Syracuse University.

Greenfield, W. D. (1982). A synopsis of research on school principals. Washington, D. C.: National Institute of Education.

Hall, G. E., Hord, S. M., Huling, L. L, Rutherford, W. L., Stiegelbauer, S. M. (1983, April). Leadership variables associated with successful school improvement. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Hall, P. M. (1983). A social construction of reality. Elementary School Journal, 84(2), 142-148.

Hall, P. M., & Spencer-Hall, D. A. (1980, June). Conditions and processes of problem identification, definition, and resolution in two school systems: Toward a grounded theory (Final report). Washington, DC: National Institute of Education (Grant NIE-G-78-0042).

BEST COPY AVAILABLE

Hanushek, E. (1970, December). The value of teachers in teaching (RM-6362-CC/RC). Santa Monica, CA: Rand.

Heyneman, S., & Loxley, W. (1983). The effect of primary school quality on academic achievement across twenty-nine high- and low-income countries. American Journal of Sociology, 88, 1162-1194.

Howell, B. (1981). Profile of the principalship. Educational Leadership, 38(4), 333-336.

Hunter, M. G. (1979). Final report of the Michigan cost-effectiveness study. Lansing: Michigan Department of Education.

Jackson, P. W. (1981a). Comprehending a well-run school: A report on a visit to a large suburban high school. Daedalus, (Special issue: America's Schools). 110(4), 81-95.

Jackson, P. W. (1981b). Secondary schooling for children of the poor. Daedalus, (Special issue: America's Schools). 110(4), 39-57.

Jencks, C. S., Smith, M., Ackland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). Inequality: A reassessment of the effect of family and schooling in America. New York: Basic Books.

Klitgaard, R. E., & Hall, G. R. (1974). Are there unusually effective schools? Journal of Human Resources, 74, 90-106.

Kounin, J. S. (1970). Discipline and group management in classrooms. New York: Holt, Rinehart & Winston.

Leiter, J. (1983). Classroom composition and achievement gains. Sociology of Education, 56, 126-132.

Levine, D. U., & Stark, J. (1981, August). Extended summary and conclusions: Institutional and organizational arrangements and processes for improving academic achievement at inner city elementary schools. Kansas City, MO: University of Missouri, Center for the Study of Metropolitan Problems in Education.

Lezotte, L. W., Edmonds, R., & Ratner, G. (1974). A final report: Remedy for school failure to equitably deliver basic school skills. East Lansing: Michigan State University, Department of Urban and Metropolitan Studies.

Lightfoot, S. L. (1978). Worlds apart: Relationships between families and schools. New York: Basic Books.

Lightfoot, S. L. (1981a). Portraits of exemplary secondary schools: Highland Park. Daedalus, (Special issue: America's Schools). 110(4), 59-80.

Lightfoot, S. L. (1981b). Portraits of exemplary secondary schools: George Washington Carver comprehensive high school. Daedalus, (Special Issue: America's Schools). 110(4), 17-37.

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

Little, J. W. (1981b, April). School success and staff development in urban desegregated schools: A summary of recently completed research. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, CA.

Little, J. W. (1981a, January). School success and staff development: The role of staff development in urban desegrated schools (Final report). Washington, DC: National Institute of Education.

Little, J. W. (1982). The effective principal. American Education, 18(7), 38-43.

Madaus, G. F., Kelleghan, T., Rakow, E. A., & King, D. J. (1979). The sensitivity of measures of school effectiveness. Harvard Educational Review, 49, 207-230.

Manassee, A. L. (1985). Improving conditions for principal effectiveness: Policy implications of research. Elementary School Journal, 85, 439-463.

Marco, G. L. (1974). A comparison of selected school effectiveness measures based on longitudinal data. Journal of Educational Measurement, 11, 225-234.

Martin, W., & Willower, D. (1981, Winter). The managerial behavior of high school principals. Educational Administration Quarterly, 17.

McCarthy, D. P., Canner, J., & Pershing, A. (1983). Local school development project: Third annual process evaluation. Office of Educational Evaluation.

McCormack-Larkin, M., & Kritek, W. J. (1983). Milwaukee's project RISE. Educational Leadership, 40, 16-21.

McLeary, L. E., & Thompson, S. D. (1979). The senior high school principalship (Volume 3: Summary report). Reston, VA: National Association of Secondary School Principals.

Metz, M. (1978). Classrooms and corridors: The crisis of authority in desegregated secondary schools. Berkeley, CA: University of California Press.

Michigan Department of Education. (n.d.). Questions and answers about Michigan Educational Assessment. Lansing: Michigan Department of Education.

Miller, J. D. (1983). Scientific literacy: A conceptual and empirical review. Daedalus, 112(2), 29-48.

BEST COPY AVAILABLE

Morris, V., & Crowson, R. (1981). The principal and instructional management. Paper prepared for Instructional Management Program. San Francisco, CA: Far West Laboratory for Educational Research and Development.

Morris, V. C., Crowson, R. L., Hurwitz, E., & Porter-Gehrie, C. (1981). The urban principal: Discretionary decision-making in a large educational organization (Report of an NIE-funded project). Chicago: University of Illinois.

Natriello, G., & Dornbusch, S. (1984). Teacher evaluative standard and student effort. New York: Longman.

New York State Department of Education. (1974a, March). Reading achievement related to educational and environmental conditions in 12 New York City elementary schools. Albany, NY: Division of Education and Evaluation.

New York State Department of Education. (1974b). School factors influencing reading achievement: A case study of two inner city schools. Albany, NY: Office of Education Performance Review. (ERIC Document Reproduction Service No. ED 089 211)

New York State Department of Education. (1976, March). Three strategies for studying the effects of school process. Albany, NY: Bureau of School Programs Evaluation.

Peterson, P., & Swing, S. (1982). Beyond time on task: Students' reports of their thought processes during classroom instruction. Elementary School Journal, 82, 481-491.

Pitner, N. (1982, March). The Mintzberg method: What have we really learned? Paper presented at the annual meeting of the American Educational Research Association, New York City.

Postlethwaite, T. H. (1975). The surveys of the International Association for the Evaluation of Educational Achievement (IEA): Implications of the IEA surveys of achievement. In A. C. Purvis and D. V. Levine (Eds.), Educational policy and international assessment. Berkeley, CA: McCutchan.

Purkey, S. C., & Smith, M. S. (1983a). Effective schools: A review. Elementary School Journal, 83(4), 427-452.

Purkey, S. C., & Smith, M. S. (1985). School reform: The district policy implications of the effective schools literature. Elementary School Journal, 85, 353,389.

Ralph, J. H., & Fennessey, J. (1983). Science or reform: Some questions about the effective schools model. Phi Delta Kappan, 64(10), 689-694.

Reynolds, D., Jones, D., & St. Leger, S. (1976). Schools do make a difference. New Society, 37, 223-225.

BEST COPY AVAILABLE

Rohrkemper, M. (1984). The influence of teacher socialization style on students' social cognition and reported interpersonal classroom behavior. Elementary School Journal, 85, 245-276.

Rowan, B., Bossert, S. T., & Dwyer, D. C. (1983). Research on effective schools: A cautionary note. Educational Researcher, 12(4), 24-31.

Rowan, B., & Denk, C. E. (1982). Modelling the academic performance of schools using longitudinal data: An analysis of school effectiveness measures and school and principal effects on school-level achievement. San Francisco, CA: Far West Laboratory for Educational Research and Development.

Rowan, B., & Miracle, A. W. Jr. (1983). Systems of ability grouping and the stratification of achievement in elementary schools. Sociology of Education, 56, 133-144.

Rowe, M. B. (1983). Scientific education: A framework for decision-makers. Daedalus, 112(2), 123-142.

Rutter, M. (1983). School effects on pupil progress: Research findings and policy. In L. S. Shulman & G. Sykes (Eds.), Handbook of teaching and policy. New York: Longman.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). Fifteen thousand hours: Secondary schools and their effects on children. Cambridge, MA: Harvard University Press.

Salley, C., McPherson, R. B., & Baker, M. E. (1979). What principals do: A preliminary occupational analysis. In D. A. Erickson and T. L. Reller (Eds.), The principal in metropolitan schools. Berkeley, CA: McCutchan.

Schwille, J., Porter, A., Belli, G., Floden, R., Freeman, D., Knappen, L., Kuhs, T., & Schmidt, W. (1983). Teachers as policy brokers in the content of elementary school mathematics. In L. S. Shulman & G. Sykes (Eds.), Handbook of teaching and policy. New York: Longman.

Shavelson, R. J. (1983). Review of research on teachers' pedagogical judgments, plans, and decisions. Elementary School Journal, 83(4), 392-413.

Shoemaker, J. (1982) What are we learning? Evaluating the Connecticut school effectiveness project. Paper presented at the annual meeting of the American Educational Research Association, New York.

Shulman, L. S. (1983). Autonomy and obligation: The remote control of teaching. In L. S. Shulman & G. Sykes (Eds.), Handbook of teaching and policy. New York: Longman.

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

BEST COPY AVAILABLE

Slavin, R. E. (1983). Realities and remedies. Elementary School Journal, 84(2), 131-138.

Soar, R., & Soar, R. (1979). Emotional climate and management. In P. Peterson & H. Walberg (Eds.), Research on teaching: Concepts, findings, and implications. Berkeley, CA: McCutchan.

Spartz, J. L., Valdes, A. L., McCormick, W. J., Myers, J., & Geppert, W. J. (1977). Delaware educational accountability system case studies: Elementary schools grades 1-4. Dover: Delaware Department of Public Instruction.

Spencer, D. A. (1984). The home and school lives of women teachers: Implications for staff development. Elementary School Journal, 84(3), 299-314.

Spencer-Hall, D. A., & Hall, P. M. (1982). Processes of problem identification and resolution in two school systems. Studies of Symbolic Interaction, 4, 191-216.

Stallings, J., & Mohlman, G. (1981, September). School policy, leadership style, teacher change and student behavior in eight schools (Final report). Washington, DC: National Institute of Education. (Grant NIE-G-80-0010).

Tomlinson, T. (1981). The troubled years: An interpretive analysis of public schooling since 1950. Phi Delta Kappan, 62, 373-376.

Travers, R. M. W. (Ed.) (1973). Second handbook of research on teaching, New York: Rand McNally.

Trisman, D. A., Waller, R. M., & Wilder, C. (1976). A descriptive and analytic study of compensatory reading programs: Final report (Vol. 2, PR 75-26). Princeton, NJ: Educational Testing Service.

Veldman, D., & Sanford, J. (1982). The influence of class ability level on student achievement and classroom behavior. Austin: University of Texas at Austin, Research and Development Center for Teacher Education.

Venezky, R. L., & Winfield, L. F. (1979). Schools that succeed beyond expectations in reading (Studies on Education Technical Report No. 1). Newark: University of Delaware. (ERIC Document No. ED 177 484)

Weber, G. (1971). Inner-city children can be taught to read: Four successful schools. Washington, DC: Council for Basic Education.

Weinstein, R. (1983). Student perceptions of schooling. Elementary School Journal, 83, 287-312.

Wellisch, J. B., MacQueen, A. H., Carriere, R. A., & Duck, C. A. (1978). School management and organization in successful schools. Sociology of Education, 51, 211-226.

BEST COPY AVAILABLE

Wittrock, M. C. (Ed.) (in press). Handbook of research on teaching (3rd ed.), New York: Macmillan.

Wolcott, H. (1973). The man in the principal's office. New York: Holt, Rinehart & Winston.