DOCUMENT RESUME

ED 333 549 EA 023 049

AUTHOR Gottfredson, Denise C.; And Others
TITLE Managing Adolescent Behavior: A Multi-Year, Multi-School Experiment. Report No. 50.
INSTITUTION Center for Research on Elementary and Middle Schools, Baltimore, MD.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Nov 90
CONTRACT OERI-G-90006
NOTE 58p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Administrator Role; Classroom Techniques; *Discipline Policy; *Discipline Problems; Educational Environment; Intermediate Grades; Junior High Schools; *Middle Schools; Parent Participation; Positive Reinforcement; Program Descriptions; Program Effectiveness; School Policy; *Student Behavior; Student Motivation; Student School Relationship
IDENTIFIERS *Charleston County School District SC

ABSTRACT

A 3-year program was implemented in eight middle schools in the Charleston County School District (South Carolina) in fall 1986 to cope with high rates of student misconduct. The program sought to improve the clarity of school rules and the consistency of rule enforcement, classroom organization and management, frequency of communication to the home regarding student behavior, and reinforcement for appropriate behavior. The use of school improvement teams and feedback to foster program implementation is described along with numerous measures used to assess program implementation. Program results indicate that the strength and fidelity of implementation varied considerably from school to school and were tied to the level of administrator support for the program. Student conduct improved significantly in schools where the program was implemented well. Speculations about effective discipline programs conclude the document. Appended are 10 tables and 1 figure. (39 references) (CLA)

Center for Research On Elementary & Middle Schools

Report No. 50
November, 1990

MANAGING ADOLESCENT BEHAVIOR: A MULTI-YEAR, MULTI-SCHOOL EXPERIMENT

Denise C. Gottfredson, Gary D. Gottfredson, and Lois G. Hybl

Center Staff

Edward L. McDill, Co-Director
James M. McPartland, Co-Director

Karl L. Alexander
Henry J. Becker
Jomills H. Braddock II
Renee B. Castaneda
Barbara S. Colton
Diane B. Diggs
Doris R. Entwisle
Joyce L. Epstein
Anna Marie Farnish
Denise C. Gottfredson
Gary D. Gottfredson
Edward J. Harsch
Brigette B. Hinte
John H. Hollifield
Lois G. Hybl
Marva J. Jeffery
Nancy L. Karweit
Melvin L. Kohn
Mary S. Leighton
Barbara M. Luebbe
Nancy A. Madden
Barbara E. McHugh
Laura B. Rice
Karen C. Salinas
Dorothy C. Sauer
Robert J. Stevens
Shi-Chang Wu

Center Liaison

Rene Gonzalez, Office of Educational Research and Improvement

National Advisory Board

Patricia A. Bauch, Catholic University of America
Jere Brophy, Michigan State University
Jeanne S. Chall, Harvard University
James S. Coleman, University of Chicago
Edgar G. Epps, University of Chicago
Barbara Heyns, New York University
Michael W. Kirst, Chair, Stanford University
Rebecca McAndrew, West Baltimore Middle School
Jeffrey Schneider, National Education Association

Managing Adolescent Behavior: A Multi-Year, Multi-School Experiment

Denise C. Gottfredson
Institute of Criminal Justice and Criminology
The University of Maryland

Gary D. Gottfredson and Lois G. Hybl

Report No. 50

November 1990

Grant No. OERI-G-90006

Published by the Center for Research on Elementary and Middle Schools, supported as a national research and development center by funds from the Office of Educational Research and Improvement, U.S. Department of Education. The opinions expressed in this publication do not necessarily reflect the position or policy of the OERI, and no official endorsement should be inferred.

Center for Research on Elementary and Middle Schools
The Johns Hopkins University
3505 North Charles Street
Baltimore, Maryland 21218

Printed and assembled by:
VSP Industries
2440 West Belvedere Avenue
Baltimore, Maryland 21215

The Center

The mission of the Center for Research on Elementary and Middle Schools is to produce useful knowledge about how elementary and middle schools can foster growth in students' learning and development, to develop and evaluate practical methods for improving the effectiveness of elementary and middle schools based on existing and new research findings, and to develop and evaluate specific strategies to help schools implement effective research-based school and classroom practices.

The Center conducts its research in three program areas: (1) Elementary Schools, (2) Middle Schools, and (3) School Improvement.

The Elementary School Program

This program works from a strong existing research base to develop, evaluate, and disseminate effective elementary school and classroom practices; synthesizes current knowledge; and analyzes survey and descriptive data to expand the knowledge base in effective elementary education.

The Middle School Program

This program's research links current knowledge about early adolescence as a stage of human development to school organization and classroom policies and practices for effective middle schools. The major task is to establish a research base to identify specific problem areas and promising practices in middle schools that will contribute to effective policy decisions and the development of effective school and classroom practices.

School Improvement Program

This program focuses on improving the organizational performance of schools in adopting and adapting innovations and developing school capacity for change.

This report, produced by the School Improvement Program, presents the findings of a three-year experiment in eight urban middle schools to test a program to improve student conduct.

Abstract

A three-year experiment in eight middle schools tested a program to improve adolescent conduct. The program sought to increase clarity of school rules and consistency of rule enforcement, increase classroom organization and management, increase the frequency of communication to the home regarding student behavior, and increase reinforcement for appropriate behavior. It was implemented in the context of an organization development approach to achieve strong implementation by increasing communication, collaboration, and planning at the school level. The strength and fidelity of implementation varied considerably across schools and was tied to the level of administrator support for the program. In schools in which the program was well-implemented student conduct improved significantly.

Author Notes

The development of this report was supported by a grant from the U.S. Department of Education, Office of Educational Research and Improvement to the Johns Hopkins University for a Center for Research on Elementary and Middle Schools. The opinions expressed do not necessarily reflect the position of the Office. We are grateful for the help of Nancy L. Karweit in designing and helping to implement the Behavior Tracking System and to Barbara Clements and Julie Sanford for providing classroom organization and management training. Renee Castañeda, Rojer Jarjoura, Laurell Wiersma, and Mark Melia provided research assistance. Many people in the Charleston County School District were partners in this research. Jeff Erickson, Larry Bolchoz, and Barbara Hess provided leadership from the district level. Karen Coste, Karl Bunch, Walter Pusey, Edward Brame, Jeff Erickson, Bob Ephriam, Barbara Cohn, Figgins Frayer, Roberta Papineau, Walter Dodson, and Nancy Gorchik led school teams. Andrea Heyer and Darlene Venable coordinated the program in the comparison schools. The help of the team members and support staff who assisted with the project is gratefully acknowledged. We also thank the teachers and students who dutifully completed surveys each quarter for three years.

Managing Adolescent Behavior:
A Multi-Year, Multi-School Experiment

Disruptive behavior in school harms both the misbehaving individual and the school community. Students who misbehave also drop out of school, use drugs and alcohol, and engage in delinquent behavior at higher rates than do their more conforming peers. Later, they tend to make poorer occupational and marital adjustments (Bachman, Green, & Wirtanen, 1971; Jessor & Jessor, 1977; Robins, 1966; Wolfgang, Figlio, & Sellin, 1972). These predictable consistencies reflect a stable pattern of antisocial behavior for some individuals, but they also suggest that school misconduct may play a part in producing negative outcomes. Suspension, a common response to school misconduct, limits students' opportunities to learn. Teachers may lower their expectations for troublesome students and limit these students' opportunities for learning by, for example, asking fewer questions. Conventional peers may avoid misbehaving students, thus forcing them to join more deviant peer groups.

Student misbehavior also has negative consequences for the school community. In a recent national survey of public school teachers, eleven percent of urban teachers mentioned fear of student reprisal as a major limitation on their ability to maintain order. Twenty-nine percent of all teachers said they seriously considered leaving teaching because of student behavior problems, and more than one-third of teachers reported that student misbehavior interferes with their teaching (National Center for Education Statistics, 1986). The consequences of fear -- teacher withdrawal, nonparticipation, alienation, and high turnover rates -- coupled with erosion of the learning environment (reductions in the amount of engaged learning time, and frustration for teachers and students) are among the prices of student misconduct.

In this article we review what is known about the nature and causes of student misbehavior, describe a middle school program designed to reduce misbehavior, and report the results of a three-year experiment to assess the effects of the program in six middle schools.

Nature of Disruptive Behavior

Behavior which results in disciplinary actions at school is best viewed as one facet of adolescent problem behavior (Jessor & Jessor, 1977). Students who lie and cheat in school also lie and cheat at home and in the community. Students who steal and fight in school often find themselves charged with criminal offenses by the police. These same students often become involved with drugs. Adolescent misconduct is usually not specialized.

Unsocialized behavior is common among adolescents (Williams & Gold, 1972). In a national sample of adolescents, 21% reported destroying others' property and 10% reported carrying a concealed weapon sometime in the last year (Elliott, Ageton, & Huizinga, 1979). In a survey of junior and senior high school students, 50% of the boys admitted hitting or threatening to hit another student (G. Gottfredson, 1987). But the conception of student misconduct as normal adolescent behavior must be tempered with the knowledge that although most adolescents engage in what might be described as "exploratory" rebellious behavior sometime during their adolescence, there exists substantial variation among individuals in the seriousness and frequency of these behaviors. Several studies have demonstrated that more than half of the officially-recorded crimes are due to a relatively small percentage of individuals (six to seven percent; Shannon, 1982; Wolfgang, Figlio, & Sellin, 1972). The distribution of disciplinary referrals across individuals in secondary schools is also lopsided: In one of the middle schools included in this study 10% of the students in the school were responsible for 45% of the office referrals during one school year.

Misbehavior in school has both individual and environmental determinants. Some environmental characteristics raise the probability of disorderly behavior in the environment and some personal characteristics make it more likely that a particular individual will misbehave.

Causes of Disruptive Behavior

Individual-level correlates. Research has identified a number of correlates of adolescent misbehavior. Misbehaving students are more likely to be male than female (Kazdin, 1987). Misbehaving youths display less academic competence, have limited career and educational objectives, dislike school, have more delinquent friends, and have lower levels of belief in conventional social rules than do their more conforming peers (G. Gottfredson, 1987; Hirschi, 1969). They also display poor interpersonal relations and are often rejected by peers because of their aggression and poor social skills. They are less likely to defer to adult authority and be polite in their interactions with adults, and they are deficient in problem-solving skills such as identifying alternative solutions to problems and taking the perspective of others (Kazdin, 1987).

Studies which have followed individuals from preschool and first grade through adolescence show that teachers' ratings of classroom disturbance, impatience, disrespect and defiance (Kellam & Brown, 1982) and "ego undercontrol" (inability to defer gratification, emotionality, bring easily irritated and angered; Block & Block, 1988) are predictive of misbehavior and psychological problems during adolescence. Clearly, a considerable amount of the disorderly behavior that occurs in schools is due to persistent and troubling characteristics of certain individuals.

Classroom-level correlates. Classroom organization and management practices also influence behavior. Disorderly behavior occurs most frequently in the absence of clearly defined classroom activities which constrain and structure student behavior (Doyle, 1986). The type of activity, physical characteristics of the setting, and level of familiarity of the student work also affect the level of disorder a classroom experiences. For example, behavior during teacher-led small group discussions is generally more orderly than behavior during independent seatwork. Loosely structured lessons, open-space physical arrangements, and higher levels of student choice and mobility all are related to higher levels of disorder. Routine tasks such as spelling tests and worksheets result in less disorder than do more complex tasks such as word problems and essays.

Evertson and Emmer (1982) highlighted the importance of teachers' organizational skills in maintaining order. Their observations of effective and ineffective classroom teachers revealed differences in the clarity and effectiveness of communication of rules and procedures, monitoring and responding to student behavior, extent of student responsibility and accountability for work, effectiveness of communication, and the organization of instruction.

School-level correlates. Discipline is also related to school characteristics. Disorderly schools are characterized by (a) teachers with punitive attitudes; (b) rules that are not perceived as fair and clear and are not firmly enforced; (c) ambiguous responses to student misbehavior; (d) disagreement among teachers and administrators about the rules and appropriate response to misbehavior; (e) students with low levels of belief in conventional social rules; and (f) a lack resources needed for teaching (G. Gottfredson & D. Gottfredson, 1985). These school characteristics are related to school disorder even when adjustments are made for characteristics of the community in which the school is located -- urbanicity, racial composition, socioeconomic status, crime level. Duke (1989) also found that school orderliness is related to the presence of a clear focus on appropriate student behavior; clear expectations for behavior; much communication about the rules, sanctions, and procedures to be used; formal discipline codes and classroom management plans; and expressed concern for students as individuals.

In summary, misbehavior in school has multiple levels of determinants. Some individuals are more likely than others to misbehave. Some teachers are more likely than others to produce higher levels of misconduct in their classrooms by their management and organization practices. Some schools more often than others fail to control student behavior.

The Present Study

The program described in this article included school-, classroom-, and individual-level interventions aimed at reducing the misbehavior of middle school students. It attempted to tighten school policies and procedures to make school

rules clearer, fairer, and more consistently enforced; establish structures and processes to promote clear communication and engage faculties in collaborative problem solving, planning, and action for school improvement; focus the goals, mission, and normative climate of the school on prosocial behavior by consistently rewarding appropriate behavior; improve classroom organization and management; and provide special services to reduce misbehavior among those most at-risk.

The Program

Context

The program was implemented in the Charleston County School District (CCSD) between Fall, 1986, and Spring, 1989. The educational and political climate at the time reflected a "get tough" approach to education. Like many urban school districts during the 1980's, standards for promotion from grade to grade were increased as part of a state-wide educational reform. Increased standards immediately increased student grade retention rates and resulted in an accumulation of "overaged" students in Charleston's middle schools. The percentage of CCSD's eighth grade students who were at least one year behind grade level increased from 34 percent in 1983 to 48 percent in 1988. As the age of the population and the history of academic failure increased, apparently so also did the level of student disattachment and misconduct. The discipline program described in this paper was an attempt to cope with the crisis of student misconduct that was produced by the accumulation in the middle schools of large numbers of students with diminished investment in education.

The Problem

The suspension rate in the CCSD middle schools rose from 41 to 100 suspensions per hundred students between 1981 and 1986, the year this project began. This high overall rate masked large differences from school to school.

For example, during the 1986-87 school year, two Charleston middle schools with similar student populations had rates of 2 and 100 suspensions per hundred students.

High rates of out-of-school suspension translate into many lost instructional days. Data from the 1987-88 school year in the six intervention schools showed that 2,042 suspensions resulted in approximately 3,850 student-instructional days lost to out-of-school suspension. In-school suspensions increased the number of lost instructional days to 7,932 during the same school year. This translated into 44 lost student years in one academic year. An examination of the reasons for a sample of suspensions in one school showed that 32% were for serious offenses (mostly fighting), 33% were attendance-related, and 35% for classroom disruption or disrespectful behavior in class (D. Gottfredson, Karweit, & G. Gottfredson, 1989).

Evidence from the Effective School Battery (G. Gottfredson, 1984b) student survey, which was administered to all middle school students during the baseline year of the study, also indicated high levels of suspension and other forms of punishment for misbehavior. Each of the eight schools surveyed scored higher than the average score for schools included in the norming sample for this battery (mostly urban secondary schools in the U.S.) on a measure which asks students to report the frequency of punishment received in school. Only one of the eight Charleston schools scored within one standard deviation of the mean, four were more than two standard deviations above the mean, and two were at the 99th percentile.

Program Components

The program was designed in collaboration with educational administrators and teachers in the Charleston County School District. A brief description of the components of the intervention strategy follows.

School Discipline Policy Review and Revision. The first component of the program called for revising the discipline policy to increase rule clarity, specifying

the consequences for specific infractions, and achieving consistency in a school-wide policy with individual classroom policies. Revised discipline policies also provided for systematically *rewarding* desired student behavior.

Behavior Tracking System. Programs that involve parents in providing consequences in the home for student behavior in school have proven effective for reducing undesirable behavior. Home-based reinforcement programs (Atkeson & Forehand, 1979; Barth, 1979; Bailey, Wolf & Phillips, 1970), which encourage parents to provide reinforcers in response to positive school behavior are effective for increasing desirable behavior. One element of home-based reinforcement involves frequent communication between the school and the home to inform the parents about the students' behavior, an activity that many schools find difficult to accomplish systematically and frequently. An objective of the program was therefore to increase the frequency of communication with parents about student behavior in school and to ensure that parents learn about positive as well as negative behavior.

A computerized Behavior Tracking System (BTS) stored information about every positive and negative referral to the office. It was used to record referral information, generate letters to the home, inform parents about positive and negative referrals to the office and about disciplinary actions taken, and generate reports for managing school discipline (e.g., detention lists, lists of students and teachers with more than a specified number of referrals, summary reports of suspensions). The system was developed to promote consistency in rule enforcement by reminding the administrator of the administrative responses allowable for each offense, according to the school's discipline code.

Classroom Organization and Management. This third component of the Charleston program was designed to replicate as closely as possible the intervention used by Texas researchers Emmer et al. (1984) which demonstrated a reduction in classroom disorder using a teacher training intervention focusing on the teacher behaviors described above. The Emmer et al. materials were used, and two of the Texas researchers provided the training. The Texas system of classroom observations for monitoring implementation of classroom practices was adapted for use in the present program.

Positive Reinforcement. Most effective classroom management strategies are built around learning principles: Consequences or events that follow a behavior affect future behavior. Reinforcers increase the behavior, punishments decrease the behavior. The research supporting the efficacy of a variety of behavioral strategies (summarized by D. Gottfredson, Karweit, & G. Gottfredson, 1989) is compelling.

The fourth component of the behavior management program was based on the assumption that misbehavior results in part because the environment reinforces undesirable behaviors and fails to reinforce desirable behaviors. It was designed to help educators structure the school environment so that (a) expectations for student behavior were understood by students and staff; (b) consequences for misbehavior were understood by students and staff; (c) misbehavior was responded to consistently and in accordance with well-communicated rules and consequences; and (d) desirable behavior was reinforced. A training manual was developed and used in conjunction with an existing book on modifying classroom behavior (Buckley & Walker, 1978). The Buckley and Walker book covered general principles of behavior modification, and the manual covered specific strategies for (a) increasing an individual's desirable behavior (contingency contracting, home-based reinforcement, and token economy), (b) decreasing undesirable behaviors (e.g., extinction, time-out, and response-cost), and (c) increasing desirable behaviors for an entire class (e.g., the "Good Behavior Game" and whole-class token economies). Teachers were expected to incorporate these behavioral practices into their instruction. During the second year of the program, students who were identified as at especially high-risk for behavior problems were targeted for special assistance using behavioral techniques.

Method of Implementation

Field research often fails because the intended interventions are not implemented as anticipated. A simplified form of the Program Development Evaluation (PDE) method (G.Gottfredson, 1984a; G. Gottfredson, Rickert, D. Gottfredson & Advani, 1984) was used to facilitate organizational change and increase implementation of the interventions. Specifically, we sought to (a) ensure

that the goals of the project and its theoretical rationale were clearly understood; (b) ensure that the objectives of the program (e.g., an increase in the perceived clarity of the school and classroom rules) were clearly understood; (c) measure the goals and objectives frequently and provide timely feedback about the extent to which goals and objectives were being met; (d) establish clear performance standards; (e) assess organizational obstacles and develop plans to overcome them; (f) monitor performance on an ongoing basis and provide workers with feedback about their performance; and (g) clearly delineate each person's responsibilities.

School Improvement Teams. A team of teachers and administrators in each school was expected to use components of the PDE method to increase the strength and fidelity of program implementation. Each principal identified five to ten team members during the 1986-87 school year (the planning year). Two members of each team (one administrator and one teacher) received brief training during that year. These team members were expected to lead their teams through a planning process to prepare the school for the program which would begin the next Fall. This group of twelve (two persons from each program school) worked together to specify concrete performance standards for each program component, and team leaders worked with their individual teams to specify standards for their specific schools. Specifying performance standards is an important step in the PDE method. These standards are expected to provide concrete, observable indicators of implementation.

The school teams also reviewed and revised their school discipline policies, oriented their faculties to the program, and developed strategies for school-wide implementation of the new practices. In each school, six to ten classroom teachers volunteered to join the school improvement teams and to become part of a staff development effort in the schools. Team members attended a training workshop to learn about the classroom organization and management and behavior change strategies. They then organized and carried out staff development workshops covering these strategies for teachers in their schools, monitored implementation of the new strategies in their colleagues' classrooms, and provided constructive feedback and ongoing technical support to their colleagues as they implemented the new practices. Two members of each school team attended quarterly meetings throughout the duration of the project to share their experiences with team

members from other schools, receive formal feedback, and identify and solve problems.

The team members were given the option of receiving a small stipend ($100) or graduate credit through a local college.

Some schools elected to rotate different faculty members onto the team in the second implementation year. The renewed teams received the same training at the beginning of the second program year that the original team members received.

Information Feedback. As a mechanism to foster implementation, feedback to individual teachers occurred informally when team members provided assistance and suggestions about ways to improve classroom practices. Teachers also received individualized feedback twice each year from student and teacher surveys (to be described below), and they received feedback about their disciplinary referrals at the discretion of the assistant principal or principal.

During the first implementation year, all teachers completed logs to record which behavior change strategies they used with their students and classes. These logs were used by team members to monitor teacher use of the new strategies. Team members also experimented with classroom observations of other teachers. They used observation forms designed to measure student engagement rates and the extent to which teachers were achieving the program's standards for classroom organization and management. These observations were dropped after the first year when efforts to overcome scheduling problems and teacher objections to what they perceived to be "evaluation" by other teachers proved unsuccessful.

Teams also received school-level feedback. Classroom climate inventories designed to measure classroom organization and management, rule clarity, teacher support, and disruptive behavior in the classroom were administered to all teachers and students each quarter. School averages for all teachers in each school were reported to the school teams four times per year. A comprehensive school assessment battery (The Effective School Battery; G. Gottfredson, 1984b) was also administered to teachers and students annually. This battery provided much of the

data necessary for the evaluation of the project, and served as an important source of information for organization development. Once per year the ESB results were presented to the administrators of the participating schools, and they were assisted in interpreting the results and considering strategies for addressing needs identified by the surveys.

Disciplinary incidents, punishments, and rewards were recorded in the computerized BTS. School teams received quarterly reports on the number and nature of disciplinary incidents and days lost to suspension for each quarter compared with the same quarter from previous years. The BTS also provided schools with the capability to generate on-the-spot summaries of referrals by teacher or student.

To summarize, the program had four components: (a) school discipline policy review and revision, (b) computerized behavior tracking, (c) improved classroom organization and management, and (d) positive reinforcement. These components were selected because prior research recommended them and because they provided a mix of activities that targeted the entire school, classrooms, and individuals within the school. These "technological" components were implemented in the context of limited organization development activity aimed at increasing school staff commitment to and ownership of the program and providing school staff with the information, planning, and management skills needed to implement the program.

Method

Subjects

The subjects were all staff and students in eight public middle schools in Charleston, South Carolina. The researchers approached central administrators in the district with a proposal to collaborate on the development and evaluation of a middle-school program. The administrators agreed and selected eight of the district's fifteen middle schools to participate. Principals in the eight schools

agreed to participate, and then the researchers designated two of the eight schools as comparison schools. This assignment was made primarily on the basis of demographics in an attempt to ensure that the comparison group was as diverse as the treatment group.

The two comparison schools were exposed to some parts of the treatment. The same data (except for data on quarterly referrals generated by the BTS system) were collected and fed back to the administrators. The assistant principals participated in quarterly feedback and planning sessions with the treatment schools. In comparison schools teams were not formed, and teachers received no training in the behavior and classroom management strategies, although the administrators had access to all program materials.

Table 1 shows some characteristics of the eight schools as we found them in 1986. All schools served grades 6 through 8. The schools ranged from majority white to majority black and from small to medium in size, and they served diverse communities.

Insert Table 1 About Here

Measures

Several different measures were used to assess the level of implementation of the program strategies, the immediate outcomes of each intervention strategy, and the ultimate program outcomes. This section summarizes these measures.

School discipline records. The BTS provided records of all referrals to the office for positive and negative behaviors. These records were used to measure the level of implementation of the positive referral system as well as changes over time in the number of negative referrals. The BTS was not introduced into the treatment schools until the Fall of 1987 (the firs program year). Paper referrals were used for all eight schools during the 1986-87 school year, and for the two

comparison schools for the duration of the project. Official reports of suspensions were collected for all eight schools for all three years from the central attendance office.

Classroom environment surveys. A brief classroom environment assessment instrument was administered to all teachers and students in all classes (except physical education classes and classes for visually and hearing-impaired students) at the end of each of the twelve academic quarters during which the program was being planned or in operation. The period during which the survey was administered was alternated each quarter to avoid taking time away from the same subject repeatedly. The survey was administered to all students in the selected classes. The percentage of the eligible classes that turned in completed student and teacher surveys was high each quarter, ranging from and 79% to 100% for any school in any given year. The average response rates for classrooms in the treatment schools for 1987, 1988, and 1989 were 93%, 84%, and 100%. For comparison schools the same percentages were 87%, 92%, and 100%. Variation in completion rate was mostly due to administrator error.

Several items from the Classroom Environment Scale (Moos & Trickett, 1974) were combined with items we generated. Factor analysis guided the formation of the following scales:

Classroom Order (teacher report) is a sixteen-item scale of Likert-type items asking the teacher to report the extent to which students engage in a range of disorderly behaviors and the extent to which disorderly behavior disrupts the learning process. The behaviors reported range in seriousness from failing to pay attention to destroying or damaging property in the class. Classes with high scores on this scale experience more order than classes with low scores. The scale is formed by averaging the sixteen items. Its *alpha* reliability, estimated with data on all classrooms in one of the baseline year quarters, is .94.

Classroom Order (student report) is a fourteen-item scale of Likert-type items asking students to report the extent to which students engage in a range of disorderly behaviors and the extent to which disorder-

ly behavior disrupts the learning process. The behaviors range in seriousness from failing to pay attention to destroying or damaging property in the class. Classes with high scores on this scale experience more order than classes with low scores. Its *alpha* reliability is .96.

Order and Organization (student report) is a five-item scale of true-false items asking students to report about the level of organization in the class and the extent to which students are engaged by what is happening in the class. Classes with high scores are more well-organized and engaged than classes with low scores. Its *alpha* reliability is .89.

Rule Clarity (student report) is a three-item scale of true-false items asking students to report on the clarity of the classroom rules. Classes with high scores are those in which students report that there is a clear set of rules to follow, and that the teacher explains the rules and the consequences for breaking them. Its *alpha* reliability is .80.

Teacher Support (student report) is a three-item scale of true-false items asking students to report about how supportive the teacher is. Classes with high scores are those in which students report that the teacher takes a personal interest in students, goes out of his or her way to help students, and is like a friend. Its *alpha* reliability is .82.

These scales are *classroom-level* scales intended to measure changes in classroom environment. The student survey scales are formed for each classroom by averaging the classroom means for each of the items. Factor analyses and reliability assessments used classroom averages for all classrooms included in one quarterly assessment.

The Classroom Order scales are intended to measure changes in classroom orderliness, a major desired outcome of the study. The Order and Organization and Rule Clarity scales were intended to measure the effect of the classroom organization and management intervention. Although increasing teacher support was not an objective of the program, items to measure it were included because previous experience with programs which sought to increase rule clarity and

consistency of rule enforcement suggested that an unintended side effect of such programs might be a decline in students' perceptions of teacher support.

Teacher ratings. Teachers whose classrooms were included in each quarterly survey administration were asked to rate the behavior of each student in the selected class. They were asked to rate how often each student "attends to academic work (i.e., pays attention, does homework, participates in class, completes classroom assignments, is cooperative and is motivated to learn)" and "disrupts the classroom (i.e., leaves seat, makes disruptive noises, speaks without permission, talks back to the teacher, fights or argues with other students, and comes late to class)." Each student was rated on a scale from "1" (almost never) to "5" (almost always). The teacher ratings for each student were averaged for all four quarters so that each student's score is based on one (12%), or the average of two (16%), three (31%), or four (41%) different ratings in each year. Variation in the number of ratings averaged to form each student's score resulted from in- and out-migration and from excluding physical education and certain handicapped classrooms from the sample. These ratings are intended to measure student classroom conduct, a major outcome targeted by the program.

Effective School Battery. The teacher and student questionnaires of the Effective School Battery (ESB, G. Gottfredson, 1984b) were administered each Spring for the three-year duration of the project. All students and teachers in the eight schools were included in the administration. The response rates for the student surveys ranged from 87% to 100%. The average response rates for the treatment schools for the three years were 93%, 90%, and 87%. For the comparison schools these averages were 93%, 96%, and 91%. The response rates for the teacher surveys ranged from 71% to 100%. The average response rates for the treatment schools for the three years were 94%, 93%, and 95%. For the comparison schools these averages were 86%, 82%, and 94%.

Several of the ESB measures are used in this study.[1] The teacher Morale, Planning and Action, and Smooth Administration scales from the teacher survey are used as measures of general organizational health, helpful for assessing the

[1]Descriptions of the scales and information about the reliability and validity of the scales are taken from G. Gottfredson (1984b).

effect of the use of the PDE method (e.g., the teams and feedback) on the general health of the organizations. The eleven-item Morale scale includes items such as "Our problems in this school are so big that it is unrealistic to expect teachers to make much of a dent in them." The nine-item Planning and Action scale includes items such as "How often do you work on a planning committee with other teachers or administrators in your school?" and the twelve-item Smooth Administration scale includes items such as "There is little administration-teacher tension in this school." G. Gottfredson (1984b, page 47) reported *alpha* reliabilities for these scales ranging from .89 to .94.

The ESB Avoidance of Punishment scale contains four items which ask students to report how often they are punished in school because of misbehavior. This scale has an *alpha* reliability of .54. Items were added to the ESB to measure student misconduct more directly. The addendum was attached to every student survey each year, so the ESB response rates apply. The addendum items included 19 items taken from the Youth in Transition study (Bachman, 1975) asking students to report how often they engaged in a variety of misbehaviors ranging in seriousness from coming late to class to trying to hurt other people. These items are averaged to form the Rebellious Behavior scale with an *alpha* reliability of .94 (computed from the 1988 survey). Scores range from one to five, with high scores indicating more rebellious behavior.

The ESB also includes measures of several intervening variables expected to respond to the treatment. The behavioral training provided to teachers was expected to change their attitudes in the nonauthoritarian direction by teaching them that student misbehavior is learned and that students can learn to replace inappropriate with appropriate behavior. The three-item Nonauthoritarian Attitudes scale containing items such as "If a pupil uses obscene or profane language in school, it should be considered a moral offense" was intended to measure these attitudes. Its *alpha* reliability is .54.

Student ESB reports of Respect for Students, Clarity of Rules, and Fairness of Rules were expected to respond to the positive reinforcement, school discipline policy revision, and the Behavior Tracking System components (the latter because it was designed to help assistant principals respond more consistently to behavior

problems). The three-item Respect for Students scale included items such as "Teachers treat students with respect" and has an *alpha* reliability of .85. The four-item Clarity of Rules scale contains items such as "Everyone knows what the school rules are" and has an *alpha* reliability of .67. The three-item Fairness of Rules scale includes items such as "The school rules are fair" and has an *alpha* reliability of .76.

The School Rewards, Positive Self-Concept, and Attachment to School scales from the ESB student survey were intended to measure student responses to the positive reinforcement component and the positive letters to the home generated from the Behavior Tracking System. The four-item School Rewards scale includes items such as "Teachers say nice things about my classwork" and the *alpha* reliability is .56. The twelve-item Positive Self-Concept scale includes items such as "Sometimes I think I am no good at all" and its *alpha* reliability is .61, and the ten-item Attachment to School scale includes items such as "This school makes me like to learn" and its *alpha* reliability is .76.

Additional items included on the addendum to the ESB teacher surveys (in the third year only) were intended to measure the level of program implementation. Teachers were asked to rate the effectiveness of the school team in five different areas on a three-point scale. Correlations among the five items were high (ranging from .56 to .75). They were averaged to form a team effectiveness scale with values ranging from one to three, with high scores indicating greater team effectiveness. Teachers were also asked to provide a global rating of the program. This is a single-item rating with response options that ranged from "two thumbs down" to "two thumbs up." Teachers were asked to report how often they used the positive reinforcement strategies in their classes, sent positive and negative communications to the home, and used the preventive management techniques included in the organization and management component. They were also asked to report on their level of use of the program materials. The scales for these items ranged from "I never saw the book" to "I read the entire book and used it to modify my practices." All of these teacher addendum items were used as single-item measures of implementation of the various program components.

Analysis Strategy

Outcome measures. Changes in school means from the 1986-87 school year (the year during which plans were made and baseline data collected) to the 1988-89 school year were examined for all outcomes. *t*-tests are used to guide interpretations of these differences for all measures for which a within-school standard deviation could be computed. ESB measures of school climate are constructed and reported at the school level. Items are aggregated to the school level before forming scales, so only between-school variation exists. The user's manual recommends forming a confidence interval for over-time comparisons by multiplying 1.96 times the standard error of measurement for the scale reported in the manual. For ESB climate scores (reported in Tables 2 and 8) differences of two or more standard errors are shown in boldface type as an aid to guard against over-interpreting small differences. Generally, only 1986-87 and 1988-89 scores and the *t*-tests for change over the entire period are reported, although we also examined changed for the interim periods. Effect sizes are also reported on several tables. These are the differences between the 1988-89 and 1986-87 scores for a school divided by the standard deviations from the 1987 survey. For ESB climate measures, the standard deviation reported in the User's Manual is used.

Insert Table 2 About Here

Implementation measures. Measures of implementation come primarily from teacher-survey addenda administered only at the end of the program. For these measures, post-treatment comparisons between high- and medium-level implementation schools are made and *t*-statistics are reported. Changes are reported separately for comparison ("low implementation") and treatment schools and for treatment schools which implemented the program with high- and medium-levels of fidelity to the implementation standards.

Results

Level of Implementation

The activities of and support for the school improvement teams varied from school-to-school, and the variation was predictable from informal assessments of building-level support for the program during the planning year.

In one of the six schools, the program was poorly communicated to the teaching staff during the planning year. The Behavior Tracking System component was overemphasized, and the components that would require more intensive staff effort were underemphasized. Team members in this school were not fully briefed on the program or their responsibilities prior to the team training. The principal was replaced at the end of the planning year, and the assistant principal (who was primarily responsible for the program) left after the first implementation year. In this school, the program was initially met with hostility from the teaching staff when they learned about the major changes they were expected to make, and the team was unable to overcome these negative feelings. The administrative team that subsequently took over was strong and supportive, and conditions improved steadily over the two-year implementation period, but the support came too late to transform an initially resisted program into a strongly supported one.

In a second school, the principal who had agreed to participate in the program was replaced just prior to the opening of the planning school year. The new principal never became extensively involved with the program because he saw it as discipline program and delegated it to the assistant principal. The assistant principal who helped to plan the program during the planning year was replaced at the end of the first implementation year by an assistant principal who was not a strong leader for the program. The team in this school was unable to make much headway without support from the administrators.

In a third school both administrators provided only weak leadership for the program. The assistant principal was replaced at the end of the first implementation year by a stronger leader, but the new assistant principal and the team

suffered under the weak leadership of the principal through the end of the program period, when he was replaced.

The other three program schools and the two comparison schools enjoyed medium- to high- levels of support and fewer administrative changes over the three-year period. These five schools had a total of four administrator changes; the three discussed above experienced five. The changes that did occur in the remaining five schools did not affect the administrative support for the program, which was initially high and remained high to medium.

The frequent meetings to provide feedback on the strength and fidelity of program implementation and the high level of communication among the school teams and between the teams and the researchers provided ample opportunity to observe differences in the levels of implementation and enthusiasm from school to school. Differences surfaced immediately and were the topic of much discussion. Consensual agreement among the researcher-practitioner planning team implied that three distinct groups participated in the program (Hess, Mack, & D. Gottfredson, 1989). Accordingly, most results are shown separately for "low implementation" schools -- the two comparison schools which participated only in some aspects of the program, "medium-implementation" schools -- the three treatment schools which experienced implementation problems stemming from leadership difficulties, and "high-implementation" schools -- the three treatment schools which enjoyed an adequate level of administrator support.

Table 2 shows school mean scores on the three organizational health measures from the 1987 (baseline) and 1989 (end of the second year of implementation) teacher surveys. The table also shows for each school the difference between the 1987 and 1989 scores divided by the standard error of measurement (SEM) for the scale from the ESB user's manual (G. Gottfredson, 1984). This ratio is useful in avoiding the over-interpretation of small differences in these school climate measures.

Table 2 shows small declines for the low-implementation schools on the organizational health measures. Only for the Planning and Action measure for one of the two schools does the difference exceed two SEMs. One of the medium-

implementation schools declined two (or almost two) SEMs on all three measures over the three-year period. Another declined substantially on teacher Morale, and the decline in teachers' reports of Smooth Administration was not trivial. The third medium-implementation school -- the first one discussed above, which received a strong administrative team too late -- improved significantly on teacher Morale and Planning and Action. The Morale improvement came only in the third year after the new administrative team was fully in place. The Planning and Action measure improved steadily over the three-year period.

Two of the high-implementation schools improved on all three measures (two of the improvements exceeded three SEMs), and one school grew worse on all three measures, although the declines may reflect measurement error.

Table 3 reports teacher perceptions of the strength of the program separately for high- and medium-implementation schools (comparison school teachers were not asked about the program since they had not experienced it). Teachers in the three high-implementation schools were generally more favorable to the program and reported that the teams were more effective than teachers in the medium-implementation schools. Team members in the high-implementation schools were also more favorably impressed with the program than their counterparts in the medium-implementation schools. Unexpectedly, team members in the medium-implementation schools rated their own effectiveness higher than the team members in the high-implementation schools.

Insert Table 3 About Here

We examined several characteristics of the teams. The number of team members at any given time ranged from four to eleven per school, with a mean of eight. The percentage of the faculty represented on the team ranged from 6% to 23%, with a mean of 16%. The percentage of team members who remained on the team for a second year ranged from 20% to 57%, with a mean of 36%. The ratio of teachers to non-teachers on the teams ranged from 0.7 to 8, with an

average of 2.9. We were unable to detect differences between the medium- and high-implementation schools on any of these team characteristics.

High- and medium-implementation schools differed on several measures of the level of implementation of the four "technological" program components -- school discipline policy review and revision, computerized behavior tracking, classroom organization and management, and positive reinforcement.

All medium- and high-implementation schools produced and published a revised school discipline handbook that described the rules and the consequences for breaking each rule. Each statement contained provisions for positive rewards for appropriate behavior to be administered at the school level. The switch-over to the BTS occurred at different points during the first implementation year for different schools. All six treatment schools were entering their office referrals and administrative responses into the BTS by the end of the first implementation year, although two of the three medium-implementation schools entered only partial information during the second year (the research staff completed the records) and never used the system to record positive referrals. All of the high-implementation schools and one of the medium-implementation schools used the BTS to record positive referrals to the office and positive responses (such as positive communications to the home and award certificates) beginning in the first year. The ratio of positive to negative actions recorded in the BTS in these four schools ranged from .24 to .71.

Information about the level of implementation of the positive reinforcement and organization and management components comes primarily from the teacher survey addendum questions. Table 4 reports these results. The teacher reports of use of contingent reinforcement of student behavior (rows 1 through 4) were higher in the high- than in the medium-implementation schools. The differences were significant or close to significant ($p < .08$ and $p < .15$) for all items except the negative communication to the home item, which was emphasized less than were the positive responses. The differences between team members' reports in high- and medium-implementation schools is larger than for the general population of teachers.

Insert Table 4 About Here

Also, the level of implementation for team members in the high-implementation schools is consistently higher than for the general teacher population in their schools. This is not the case in the medium-implementation schools where the level of implementation is roughly equivalent for team and non-team teachers. Although the program targeted all teachers and was not intended to produce a difference between team members and other teachers in their practices, the different patterns in the high- and medium- implementation schools converge with other information about differences in level of implementation found for the high- and medium-implementation schools.

Two items on the teacher addendum asked about the level of implementation of the classroom organization and management strategies. For both the general teacher and the team member populations the reports of intervention to prevent misbehavior were higher in the high-implementation schools. Teachers in the medium-implementation schools, however, reported that they had read the Emmer et al. book as carefully as teachers in the high-implementation schools.[2]

Program Outcomes

Tables 5 through 7 report changes from baseline to the end of the program on the primary outcomes of the program -- classroom order and student behavior. Tables 5, 8, and 9 also report changes on intermediate program outcomes targeted by the program. Table 10 summarizes the effect sizes for the changes on measures included in the ESB surveys.

[2]The relatively high level of use of this book in one of the medium-implementation schools probably resulted from a principal's mandate that every teacher read the book.

Insert Tables 5, 6, 7, 8,and 9 About Here

Table 5 shows that treatment schools improved on all measures included in the quarterly classroom environment assessments, and that these improvements were highly significant for student reports of classroom order, classroom organization, and rule clarity. The low-implementation (comparison) schools declined slightly on all measures except rule clarity, which stayed the same. The breakdown by level of implementation shows that the treatment effects are found mostly fo. the high-implementation schools except for Rule Clarity, on which the medium-implementation schools also improved significantly.

Figure 1 shows the quarterly scores for student reports of classroom orderliness. The graph shows that the medium- and low-implementation schools' performance were similar: They are marked by the same pattern of seasonal variation in classroom orderliness, but there is gradual improvement, albeit slight, over the twelve quarters. The high implementation schools started off with significantly lower classroom order but ended up with the highest. Their scores are marked by the same ebb and flow pattern as the other schools, but their improvement is more dramatic over the project period.

Insert Figure 1 About Here

Table 6 shows the average teacher ratings of student behavior for 1987 and 1989. Teacher ratings of student attentiveness increased and of student disruptive behavior decreased significantly in the high-implementation schools. Teacher ratings of disruptive behavior increased significantly in the medium-implementation schools. Neither ratings changed significantly in the low-implementation schools. Effect sizes are generally small.

Table 7 shows school averages for students' reports of their own misbehavior and their schools' responses to misbehavior. In all but one school, rebellious

behavior significantly increased between 1987 and 1989. These increases were most marked in the medium implementation schools. Students' reports of the level of punishment they received in school declined significantly over the time period for all three high-implementation schools.[3] Punishment increased in all three medium-implementation schools, significantly in one. The level remained the same in one and declined significantly in another low-implementation school.

Suspensions and disciplinary referrals are often used as indicators of program success or failure in evaluations of discipline programs. We collected quarterly reports of suspensions and disciplinary referrals in each of the eight schools participating in the study. Our examination of these figures suggested that they are primarily measures of administrator discipline style. As discussed earlier, five of the six program schools experienced at least one administrator change during the project period. The suspension and disciplinary referral rates changed dramatically with these shifts.

Changes in rates of referral and suspension over the three-year period for the two comparison and one high-implementation program school which experienced no administrative changes show that the one stable program school cut its referral rate by 62% and increased its suspension rate by 43% over the project period. One comparison school increased its referral rate by 28% and its suspension rate by 18% in the same period, and the other increased its referral rate by 10% and decreased its suspension rate by 12% in the same period. The meaning of these changes is not clear, but they do show that the one stable program school changed its referral and suspension policies and procedures during the period. This change is more a reflection of the implementation of new discipline procedures and policies under the revised school discipline plan than of student behavior.

Table 3 shows changes in school means on four intermediate outcomes targeted by the program -- teacher Nonauthoritarian Attitudes, Respect for Students, Clarity of Rules, and Fairness of Rules. The Nonauthoritarian Attitudes and Respect for Students measures improved in most of the schools, but the

[3]Note that the scoring is in the positive direction such that high scores indicate fewer reports of punishment.

improvements were greatest in the high-implementation schools. The only meaningful changes on the school rules measures were increases for program schools. One of the high-implementation schools improved substantially on Clarity of Rules and on Fairness of Rules.

Changes on the intermediate student experiences and attitudes targeted by the program are shown in Table 9. Students in five of the six program schools and one of the comparison schools reported more school rewards in 1989 than in 1987. These improvements were significant in four of the schools. Student self-concepts became more positive in all of the high-implementation schools, significantly in one. This outcome measure declined in two of the three medium-implementation schools, significantly in two. Changes in the low-implementation schools were not significant. Student attachment to school declined in the medium-implementation schools, significantly in one. It improved in two of the high-implementation schools (significantly in one) but declined significantly in another. Attachment barely changed in the low-implementation schools.

Summary

In summary, variation in the level of implementation of the organization development components of the program was directly related to the administrators' support for the program at the building level. Three of the six program schools had little support. These schools' organizational health deteriorated over the course of the project, and the program was perceived as less effective than in the schools with higher levels of support. The faculties in the schools enjoying higher levels of administrator support reported somewhat higher levels of organizational health over the three-year period.

Measures of implementation of the "technological" program components suggested that the School Discipline Policy component, including provisions for school-wide positive reinforcement for appropriate behavior, was implemented in all six treatment schools. The computerized Behavior Tracking System was fully implemented in all high-implementation and one medium-implementation school. The classroom-level positive reinforcement and classroom organization and

management components were implemented more faithfully in the high- than in the medium-implementation schools.

Table 10 summarizes the mean differences found on Tables 7 through 9 for schools in different implementation groups, by showing the effect sizes for the changes. Results for the measures of disruptive behavior accord with the conclusion that the program had beneficial effects on student behavior when it was well-implemented. Students' reports of their own rebellious behavior generally increased over the three-year period, but the increase was less dramatic in the high-implementation schools than in other schools. Teachers' perceptions of their students' behaviors (from Table 6) improved in the high- and deteriorated in the medium-implementation schools. They did not change in the low-implementation schools. Students' reports of the orderliness of their classrooms (from Table 5) increased in the high-implementation schools only. Students' reports of the extent of punishment they experienced in school, a measure both of student behavior and level of implementation of the program, suggested that students were punished less frequently in the high-implementation schools, more frequently in the medium-implementation schools, and about the same in the low-implementation schools.

Insert Table 10 About Here

Respect for Students and Fairness of Rules increased for treatment as well as comparison schools, but the increases for treatment schools are due only to the high-implementation schools. Clarity of Rules increased for all treatment schools and declined for comparison schools. Authoritarian teacher attitudes declined only in the program schools. The high-implementation schools experienced large changes in Fairness of Rules, Respect for Students, and Clarity of Rules and moderate changes in Avoidance of Punishment and Nonauthoritarian Attitudes. All changes were in the desired direction, except for Rebellious Behavior. The change in Rebellious Behavior was less than one-tenth of a standard deviation. The medium-implementation schools experienced moderate improvements in the Clarity of Rules measure and small to moderate improvements in student reports of School Rewards and teacher Nonauthoritarian attitudes. Rebellious Behavior

increased substantially, and several of the student experiences and attitudes changed in the undesired direction.

Discussion and Supplementary Explorations of the Results

Some of the results summarized above require further discussion.

Are beneficial outcomes related to actual implementation of program strategies? Teachers in high-implementation schools reported using program strategies more often than teachers in medium-implementation schools (Table 4). By itself, this difference is ambiguous. Because pre-treatment measures of discipline practices were not taken, it is not possible to confidently interpret the post-treatment difference as a beneficial program effect. It may instead reflect pre-existing differences in the use of these practices. Similarly, the results for team members showing that team members in high-implementation schools used the practices more than team members in medium-implementation schools suggests either a beneficial effect on team-member practices in high-implementation schools, *or* that the high-implementation schools selected as team members teachers who were more likely to use these practices anyway.

In order to examine the plausibility of the differential selection hypothesis, we compared pretreatment Classroom Environment Assessment measures for classes of teachers who eventually became team members and those who did not. Team members' classrooms were significantly more organized and orderly, had clearer rules, and experienced more teacher support than other teachers' classrooms. But this pre-existing difference was as apparent in the medium-implementation schools as it was in the high-implementation schools. Therefore, differential selection does not appear to explain the results.

Additional evidence that team membership in the high-implementation schools had a beneficial effect on teacher practices comes from a supplementary regression analysis in which the Classroom Environment Assessment scores for the final program year were regressed on a dummy variable measuring team membership and the same Classroom Environment Assessment score from the

previous year. This analysis tells us the extent to which change from 1988 to 1989 in the classroom environment were related to team membership. Larger regression weights in the high-implementation schools than in the medium-implementation schools would support the conclusion that team membership produced changes in teacher practices which subsequently effected the classroom environment. For all schools combined -- and with the prior year's measure controlled -- team membership was hardly related to 1988-89 teacher reports of classroom orderliness (*beta* = .02, p=.66). A significant (p=.02) interaction was found for team membership with the level of implementation of the program, however. Team membership was related to increased classroom orderliness (*beta* = .30, p=.08) in the high implementation schools. In the medium-implementation schools, team membership was related to *lower* classroom orderliness (*beta* = -.18; p=.14). Results for the other measures of classroom environment followed the same pattern, i.e., with the team membership being more highly related to the outcome in the high implementation schools, but the interactions were not statistically significant. These results help to reject the alternative hypothesis that the higher level of implementation reported by teachers and team members in the high implementation schools -- and the beneficial outcomes observed for these schools -- can be explained by a selection artifact.

These results increase confidence in the interpretation that team membership in the high-implementation schools altered teacher practices and improved classroom conditions.

Why did low-implementation schools perform better than medium-implementation schools? The two low-implementation schools were intended to be no-treatment comparison schools. Practical considerations prevented them from being left entirely alone, however. They were considered to be part of the district planning team that was charged with developing, implementing, and evaluating the program. They were promised top priority in the next phase of the program, which was to extend the effective elements from the experiment into other middle schools. The comparison schools, then, were very much a part of the program. Their assistant principals attended all planning and feedback sessions and were given all materials. Their principals participated in annual feedback workshops.

Table 10 and Table 6 show that several of the outcome measures in the study improved most in the high-implementation schools and least in the medium-implementation schools. The low-implementation schools experienced more beneficial change than did the medium-implementation schools on teacher ratings of student behavior (Table 6), Rebellious Behavior, Avoidance of Punishment, Respect for Students, Fairness of Rules, Positive Self-Concept, and Attachment to School (Table 10). A closer look at the low-implementation schools helps to explain why.

For the most direct measures of student behavior, large differences exist in the pre-post changes observed for the two comparison schools. In one (School "A"), significant increases in teacher ratings for attending to work and significant declines in teacher ratings of disruptive behavior were observed. In the other (School "B"), significant changes in the undesired direction on both ratings were observed. School A's students reported more rebellious behavior over the three-year period, but (as in the high-implementation schools) the increase was not statistically significant. School B's students reported significantly more rebellious behavior (as did students in the medium-implementation schools). School A's students reported significantly less school punishment and significantly more rewards over the period. School B's students experienced more punishment and fewer rewards. Clearly, School A resembles the high-implementation schools on many of the measures that might be expected to be most responsive to changes in the consequences for misbehavior at the school, and School B resembles the medium-implementation schools. Note that differences between School A and School B were not apparent for outcomes expected to be more responsive to the organization development or classroom organization and management activities in the school.

In an attempt to discover what discipline practices at School A might have produced results similar to those in high-implementation schools, we performed informal, open-ended interviews with teachers in the three high-implementation schools and School A. Ten teachers from each school were selected on the basis of positive changes in their classes' classroom environment assessment scores. A research assistant interviewed all forty teachers about their usual discipline practices. Through these interviews we learned that the assistant principal in School

A provided a two-day discipline workshop during the 1987-88 school year (after the planning year for our project). This workshop used the Assertive Discipline (Canter & Canter, 1976) model to focus on increasing clarity and consistency of rule enforcement at the classroom level. Sixty-four percent of the teachers interviewed in School A (compared with 67%, 44%, and 11% in the other schools) mentioned Assertive Discipline as having helped them with discipline problems. Eighty-two percent (compared with 67%, 78%, and 44% in the other schools) specifically mentioned using a "checks on the board" system to introduce progressively harsher penalties for misbehavior. When asked to state the single most important factor in their success as disciplinarians, the School A teachers mentioned rule clarity and consistency of enforcement at much higher rates than the teachers in the high-implementation schools (73% vs. 33%, 56%, and 56%).

We also learned that the teachers in School A used strategies for contingent reinforcement of behavior at a high rate. Eighty-two percent of the teachers in School A reported that they apply consequences for behavior according to a prearranged plan. In the high-implementation schools, 78%, 78%, and 67% of the teachers reported frequent use of these methods, which were part of the program training. Roughly the same percentage of teachers in School A mentioned using specific behavior change methods included in the program training as teachers in the high-implementation schools (91% vs. 89%, 100% and 100%). Many of the specific methods for positive reinforcement mentioned by School A teachers were identical to those used in the high-implementation schools.

Teachers in the high-implementation schools mentioned the specific program materials, help from the school team, and feedback produced as part of the program as having been helpful. None of the teachers in School A mentioned these specific materials and support structures. They also mentioned classroom organization and management issues as important in maintaining discipline more than did teachers in School A, and they reported having special plans for improving behavior of at-risk youth more than the teachers in School A. For example, 80%, 67%, and 83% of the teachers in the high-implementation schools said they used contracts with at-risk youths. Only 50% of the School A teachers

used this program strategy. However, one teacher in School A reported that the assistant principal targeted specific high-risk individuals for special assistance and used the kind of monitoring and contracting strategies included in our program materials.

In summary, one of the low-implementation schools was clearly engaged in a major behavior management program, and the program resembled in many ways the program that was implemented in the high-implementation schools. It included a positive reinforcement component and those segments of the classroom organization and management component that overlapped with Assertive Discipline, those focusing on clarity and consistency of classroom rules. It did not include the other organization development components or the computerized behavior management system with frequent communications to the home. These differences help to explain the beneficial changes observed for School A and lend credibility to the conclusion that the program, when well-implemented, was effective for reducing behavior problems.

Why did rebellious behavior increase? All results are consistent with the interpretation that the program produced beneficial effects on student behavior. The results for Rebellious Behavior, however, are counter-intuitive because they show that these behaviors generally increased in the eight schools included in the study. The effect of the program was to limit the undesired change in the high-implementation schools.

As noted earlier, misbehavior has several sources. It is produced by school- and classroom-level factors as well as individual-level attitudes, beliefs and experiences which place individuals at higher risk for engaging in misbehavior. These individual characteristics include low levels of attachment to school, low commitment to the pursuit of educational and occupational goals, and low beliefs in and internalization of moral rules (Hirschi, 1969). We discussed earlier the general educational climate in Charleston during the project period, and suggested that the school system's control over its clientele waned during the period. This climate of low social control allowed rebellious and delinquent behavior to flourish, as reflected in the increasing scores on rebellious behavior. This interpretation fits with the pattern of results that showed increases in the most

direct measures of student misconduct -- self-reports of rebellious behavior and teacher ratings of student behavior -- and little change over time on other measures of misconduct such as classroom orderliness, which is more sensitive to the classroom- and school-level sources of misbehavior.

Conclusions

Emmer and Aussiker (1989) recently reviewed research on four popular discipline programs. They found some evidence of positive effects on teacher attitudes, beliefs, and perceptions, mixed evidence for effects on teacher behaviors, and almost no evidence supporting a positive effect on student behavior. The work summarized here implies that schools can intervene to improve student behavior. The features which distinguish this moderately effective program from those reviewed by Emmer and Aussiker are (a) researchers and practitioners collaborated in the development and implementation of the program, (b) the technological components of the program were supported by organization development activities and structures, and (c) the program attempted to target sources of misbehavior from multiple levels.

Enthusiasm for this approach to improving student conduct in the middle grades must be tempered. First, building-level administrative support for discipline programs is an essential element. Discipline is at the heart of the instructional program because student misconduct directly effects instructional effectiveness and learning. Disciplinary practices are an essential element of the instructional program and must be coordinated, monitored and improved just as any other instructional practice must be. When the school administration gives low priority to classroom and behavior management by failing to create a structure to promote professional growth in this area, the teachers also may give discipline low priority. Additional research and development is needed to learn how to foster organizational competence to support change. We now assume that much effort is wasted when training programs are not augmented with school-level support structures to facilitate change.

Other features of organizational health also facilitate strong program implementation. Results from this experiment accord with results from earlier experiments (D. Gottfredson, 1987) to suggest that decision-making structures which allow the workers with direct responsibility for implementing new practices to participate in planning and refining those practices are effective for raising morale and for directing energy towards change. Healthy organizations are more capable of planning for and managing innovation, and they produce stronger programs.

The classrooms of those teachers who participated as members of a school improvement team that was responsible for providing leadership for the program and for tailoring the program to fit their school's unique needs had more positive outcomes (at least in the schools in which the team received strong administrator backing). These teachers received more training in the program components than the other teachers in the school, but the study results suggest that it was the team participation rather than the additional training which led to the improved outcomes. Teachers in the schools whose teams were ineffective also received more training, but their classes did not improve. This outcome underscores the importance of organizational support for professional development as opposed to mere training, a point that has been emphasized by others as well (Berman, 1981; Miles, 1981).

Some Speculations

The study provides some insights into the design of effective discipline programs. All of the schools in the study clarified their school discipline policies, implemented a computerized behavior management system, and developed a school-level system for providing rewards for appropriate behavior. These changes were evident in the increases in students' reports of clarity of rules and rewards. These changes alone were not sufficient to reduce student misconduct and rebellious behavior. Only those schools which significantly reduced the amount of punishment concurrently with the other changes experienced beneficial student outcomes. This suggests that simply adding a cosmetic system of positive

reinforcement onto a punitive system is not productive. The underlying approach to discipline must be examined to produce a coherent system.

Schools which implemented only the school-level components did not experience positive change. Most misbehavior can probably be traced to classroom- and individual-level sources which the school-level components did not alter. While the school-level components probably helped to set the context for alterations in these more important domains, they were unable to stand on their own.

Classroom-level changes, on the other hand, appeared effective for modifying student behavior. We are unable to disentangle the effects of preventive classroom management changes from changes that targeted troublesome students using the behavior change strategies provided (e.g., contracting), but it is clear that a combination of these approaches was effective for producing improvements both in the orderliness of the classroom environment and in teacher and student reports of student behavior.

Future work in this area should determine the relative effectiveness of approaches to behavior change that target individual students at elevated risk for behavior problems and those that alter the classroom and school environment to prevent misbehavior. Although recent statements on the future of discipline programs seem to favor classroom environmental approaches (Doyle, 1986), we lack strong tests of the relative efficacy of *plausible* targeted programs on the one hand and environmental programs on the other. Targeted programs which use behavioral and cognitive approaches to teach students how to manage their own behavior appear highly effective for replacing inappropriate with appropriate behavior (Alexander & Parsons, 1982; Barth, 1979; Manning, 1988; Patterson, Chamberlain, & Reid, 1982; Schinke & Gilchrist, 1984; Spivack, Platt, & Shure, 1976). In all likelihood, a behavior management program that combined these theoretically plausible targeted approaches with environmental change approaches would prove most beneficial.

References

Alexander, J. F., & Parsons, B. V. (1973). Short-term behavioral intervention with delinquent families: Impacts on family process and recidivism. *Journal of Abnormal Psychology, 81,* 219-225.

Atkeson, B. M. & Forehand, R. (1979). Home-based reinforcement programs designed to modify classroom behavior: A review and methodological evaluation. *Psychological Bulletin, 86,* 1298-1308.

Bachman, J. G. (1975). *Youth in transition (ICPSR 3505) documentation manual, Volume 1.* Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Bachman, J. G., Green, S., & Wirtanen, I. D. (1971). *Youth in Transition, Vol. 3: Dropping Out - Problem or Symptom?* Ann Arbor, MI: University of Michigan Institute for Social Research.

Bailey, J. W., Wolf, M. M., & Phillips, E. L. (1970). Home-based reinforcement and the modification of pre-delinquent classroom behavior. *Journal of Applied Behavior Analysis, 3,* 223-233.

Barth, R. (1979). Home-based reinforcement of school behavior: A review and analysis. *Review of Educational Research, 49,* 436-458.

Berman, P. W. (1981). Educational change: An implementation paradigm. In R. Lehming & M. Kane (Eds.), *Improving schools: Using what we know.* Beverly Hills, CA: Sage.

Block, J., & Block J. (1988). Longitudinally foretelling drug usage adolescent: Early childhood personality and environmental precursors. *Child Development, 59,* 336-355.

Buckley, N. K., & Walker, H. M. (1978). *Modifying classroom behavior.* Champaign, IL: Research Press.

Canter, L. & Canter, M. (1976). *Assertive Discipline.* Los Angeles, CA: Lee Canter & Associates, Inc.

Doyle, W. (1986). Classroom organization and management. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd edition). New York: Macmillan.

Duke, D. L. (1989). School organization, leadership, and student behavior. In Moles, O. C., *Strategies to reduce student misbehavior*. Washington, DC: U.S. Department of Education.

Elliott, D. S., Ageton, S. S., & Huizinga, D. (1979). *1978 self-reported delinquency estimates by sex, race, class and age* (Report No. 9). Boulder, CO: Behavioral Research Institute.

Emmer, E. T., & Aussiker, A. (1989). School and classroom discipline programs: How well do they work? In Moles, O. C., *Strategies to reduce student misbehavior*. Washington, DC: U.S. Department of Education.

Emmer, E. T., Evertson, C. M., Sanford, J. P., Clements, B. S., & Worsham, M. E. (1984). *Classroom management for secondary teachers*. Englewood Cliffs, NJ: Prentice-Hall.

Evertson, C. M., & Emmer, E. T. (1982). Effective management at the beginning of the school year in junior high classes. *Journal of Educational Psychology, 74*, 485-498.

Gottfredson, D. C. (1987). An evaluation of an organization development approach to reducing school disorder. *Evaluation Review, 11*, 739-763.

Gottfredson, D. C., Karweit, N. L., & Gottfredson, G. D. (1989). *Reducing disorderly behavior in middle schools* (Report No. 37). Baltimore: The Johns Hopkins University, Center for Research on Elementary and Middle Schools.

Gottfredson, G. D. (1984a). A theory-ridden approach to program evaluation: A method for stimulating researcher-implementer collaboration. *American Psychologist, 39*, 1101-1112.

Gottfredson, G. D. (1984b). *Effective School Battery: User's manual*. Odessa, FL: Psychological Assessment Resources.

Gottfredson, G. D. (1987). American education--American delinquency. *Today's Delinquent. 6*, 5-70.

Gottfredson, G. D., & Gottfredson, D. C. (1985). *Victimization in schools*. New York: Plenum.

Gottfredson, G. D., Rickert, D. E., Gottfredson, D. C., & Advani, N. (1984). Standards for Program Development Evaluation plans. *Psychological Documents, 14*, 32. (Ms. No. 2668)

Hess, B., Mack, D., & Gottfredson, D. (1989). *Interim report on Project BASIS*. Unpublished manuscript. Baltimore: Johns Hopkins University, Center for Research on Elementary and Middle Schools.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley, CA: University of California Press.

Jessor R., & Jessor S. L. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth*. Boulder, CO: University of Colorado Institute of Behavior Science.

Kazdin, A. E. (1987). *Conduct disorders in childhood and adolescence*. Newbury Park, CA: Sage.

Kellam, S. G., & Brown, C. H. (1982). *Social adaptational and psychological antecedents in first grade of adolescent psychopathology ten years later*. Paper presented at the Research Workshop on Preventive Aspects of Suicide and Affective Disorders Among Adolescents and Young Adults, Harvard School of Public Health, Boston, MA.

Manning, B. H. (1988). Application of cognitive behavior modification: First and third graders' self-management of classroom behaviors. *American Educational Research Journal, 25*, 193-212.

Miles, M. (1981). Mapping the common properties of schools. In R. Lehming & M. Kane (Eds.), *Improving Schools: Using What We Know*. Beverly Hills, CA: Sage.

Moos, R. H. & Trickett, E. J. (1987). *Manual for the Classroom Environment Scale*. Palo Alto: Consulting Psychologists Press.

National Center for Education Statistics. (1986). *Discipline in public secondary schools*. Washington, DC: U.S. Department of Education.

Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of parent training procedures. *Behavior Therapy, 13*, 638-649.

Robins, L. N. (1966). *Deviant children grown up*. Baltimore: Williams & Wilkins.

Shannon, L. W. (1982). *Assessing the relationship of adult criminal careers to juvenile careers*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Schinke, S. P., & Gilchrist, L. D. (1984). *Life skills counseling with adolescents*. Austin, TX: Pro-Ed.

Spivack, G., Platt, J. J., & Shure, M. B. (1976). *The problem-solving approach to adjustment*. San Francisco: Jossey-Bass.

Williams, J. R. & Gold, M. (1972). From delinquent behavior to official delinquency. *Social Problems, 20*, Fall, 209-29.

Wolfgang, M. E., Figlio, R., & Sellin, T. (1972). *Delinquency in a birth cohort*. Chicago: University of Chicago Press.

Table 1

Selected Characteristics of the Participating Schools

School number	Total enrollment	Percentage white	Affluence index[a]	Location[b]
Treatment:				
1	554	73	1.38	S
2	944	51	1.87	M
3	781	36	2.07	M
4	724	49	2.26	M
5	1050	50	1.65	S
6	460	3	2.51	U
Comparison:				
1	490	70	1.35	S
2	716	0	2.69	U

[a] Affluence index is scored "1" if student does not receive free or reduced lunch cost, "2" for reduced lunch cost, and "3" for free lunch. Index is based on all students in grades 6 and 8 in Spring, 1987.

[b] S=suburban; M=mixed (urban and suburban); U=urban

Table 2

Teacher Reports of School Organizational Health, 1987 and 1989, by Level of Implementation

Level of implementation and school	Measure of organizational health: Teacher Morale 1987	Teacher Morale 1989	Teacher Morale Diff/ St. err.	Planning and Action 1987	Planning and Action 1989	Planning and Action Diff/ St. err.	Smooth Administration 1987	Smooth Administration 1989	Smooth Administration Diff/ St. err
Low (comparison)									
School 1	1.69	1.64	-1.24	1.69	1.61	-2.01	1.72	1.69	- .70
School 2	1.78	1.72	-1.34	1.63	1.60	- .90	1.70	1.71	.03
Average			-1.29			-1.46			- .34
Medium									
School 3	1.51	1.64	**2.78**	1.55	1.75	**4.68**	1.74	1.77	.68
School 4	1.60	1.46	**-3.06**	1.56	1.57	.15	1.81	1.75	**-1.43**
School 5	1.72	1.59	**-2.98**	1.73	1.65	**-1.97**	1.80	1.68	**-2.82**
Average			-1.09			.95			-1.19
High									
School 6	1.73	1.66	**-1.45**	1.68	1.66	- .54	1.84	1.79	**-1.04**
School 7	1.76	1.84	**1.74**	1.62	1.64	.54	1.63	1.77	**3.25**
School 8	1.64	1.80	**3.49**	1.71	1.75	.92	1.79	1.83	.98
Average			1.26			.31			1.06

Note. **Bold entries indicate that the difference between the 1987 and 1989 climate scores exceeds twice the standard error of measurement.**

Table 3

Means and Standard Deviations on Teacher Reports of Team Effectiveness and Overall Satisfaction with Program--by Level of Implementation and Team Membership

Teacher report	Level of implementation					
	High			Medium		
	M	SD	N	M	SD	N
Percentage "thumbs up"						
All teachers	.59**	.49	98	.33	.47	110
Team members	.86**	.36	14	.42	.50	26
Team effectiveness						
All teachers	2.32**	.58	98	2.05	.62	107
Team members	1.38*	.45	14	1.78	.65	24

* Difference between high- and medium-implementation schools is significant at the $p<.05$ level.

**Difference between high- and medium-implementation schools is significant at the $p<.01$ level.

Table 4

Means and Standard Deviations on Measures of Program Implementation, by Level of Implementation and Team Membership

Implementation measure	Level of implementation: High M	High SD	High N	Medium M	Medium SD	Medium N
Teacher reports: All teachers						
Rewards for individual behavior	4.08	.80	105	3.89	.82	113
Rewards for group behavior	3.88**	.94	104	3.44	.92	113
Negative communication to home	3.82	.78	105	3.70	.76	115
Positive communication to home	3.80	.75	105	3.65	.81	114
Read contingency management book	2.14	1.14	86	2.06	1.02	102
Intervenes to prevent misbehavior	4.02**	.76	105	3.74	.62	115
Read organization and management book	2.16	1.12	86	2.18	1.08	104
Teacher reports: Team members only						
Rewards for individual behavior	4.36**	.63	14	3.77	.59	26
Rewards for group behavior	4.28**	.73	14	3.46	.71	26
Negative communication to home	4.14**a	.86	14	3.56	.75	27
Positive communication to home	4.07*	.83	14	3.48	.89	27
Read contingency management book	3.08	.95	13	2.68	1.11	25
Intervenes to prevent misbehavior	4.28**	.61	14	3.74	.66	27
Read organization and management book	3.15	.99	13	2.68	1.14	25

* Difference between high and medium implementation schools is significant at the $p<.05$ level.

**Difference between high and medium implementation schools is significant at the $p<.01$ level.

a Team membership by school level of implementation interaction is significant at the $p<.05$ level.

Table 5

Means and Standard Deviations for Classroom Environment Assessment Measures, Three Years, by Level of Implementation

Measure	All treatment			High implementation			Medium implementation			Low implementation (comparison)		
	M	SD	N	M	SD	N	M	SD	N	M	SD	N
1987 (baseline)												
Classroom Order, Students	3.28	.49	211	3.20	.52	103	3.36	.44	107	3.44	.39	59
Classroom Order, Teachers	3.71	.61	211	3.72	.66	103	3.71	.55	107	3.90	.50	59
Classroom Organization	1.67	.16	211	1.65	.17	103	1.69	.14	107	1.73	.14	59
Rule Clarity	1.83	.10	211	1.81	.11	103	1.84	.09	107	1.85	.07	59
Teacher Support	1.66	.16	211	1.63	.17	103	1.68	.16	107	1.68	.16	59
1988												
Classroom Order, Students	3.38	.51	206	3.42*	.52	99	3.37	.50	101	3.45	.43	61
Classroom Order, Teachers	3.68	.62	207	3.70	.64	97	3.67	.60	103	3.85	.47	60
Classroom Organization	1.70	.16	207	1.70	.17	99	1.70	.15	102	1.73	.15	61
Rule Clarity	1.86*	.09	207	1.86*	.10	99	1.87*	.07	102	1.86	.10	61
Teacher Support	1.68	.16	206	1.67	.16	99	1.69	.15	101	1.66	.17	61
1989												
Classroom Order, Students	3.45**	.52	226	3.52**	.50	117	3.37	.53	109	3.43	.45	62
Classroom Order, Teachers	3.75	.58	229	3.79	.61	117	3.69	.55	109	3.85	.46	61
Classroom Organization	1.72**	.16	226	1.74**	.16	117	1.70	.17	109	1.70	.15	62
Rule Clarity	1.87**	.10	226	1.87*	.11	117	1.87*	.10	109	1.85	.09	62
Teacher Support	1.68	.15	226	1.69*	.14	117	1.67	.16	109	1.65	.15	62

Note. Means are averages for all classrooms participating in the survey for all quarters each year.

* Difference between mean for starred year and baseline year is significant at the $p<.05$ level.
**Difference between mean for starred year and baseline year is significant at the $p<.01$ level.

Table 6

Change from 1987 to 1989 in Teacher Ratings of Students' On-task Behavior, by Level of Implementation

Behavior and group	1987			1989			Effect size
	M	SD	N	M	SD	N	
Attends to academic work[a]							
All treatment	3.87	.93	4011	3.88	.96	4066	.01
High implementation	3.91*	.89	1951	3.99	.91	1973	.09
Medium implementation	3.82	.96	2060	3.77	1.00	2093	-.05
Low implementation (comparison)	3.87	.88	1162	3.90	.83	1215	.03
Disrupts the classroom[a]							
All treatment	2.00	.96	4006	2.01	.98	4064	.01
High implementation	2.02**	.94	1952	1.91	.92	1977	-.12
Medium implementation	1.98**	.97	2054	2.10	1.02	2087	.12
Low implementation (comparison)	1.95	.90	1160	1.95	.93	1214	.00

* Change from 1987 to 1989 is significant at the $p<.05$ level.
**Change from 1987 to 1989 is significant at the $p<.01$ level.

[a] Coded "5" = almost always; "1" = almost never.

Table 7

School Means for Measures of Student Misbehavior, 1987 and 1989

Level of implementation and school	Rebellious Behavior		Avoidance of Punishment	
	1987	1989	1987	1989
Low (comparison)				
School 1	.91	1.09	.71	.71
School 2	.83	.90*	.62	.65*
Medium				
School 3	.81	1.16**	.73	.69*
School 4	.87	1.10**	.78	.77
School 5	.86	1.11**	.74	.72
High				
School 6	.87	.94*	.76	.79*
School 7	.96	1.08	.70	.75**
School 8	.96	.95	.63	.70**

* Difference between 1987 and 1989 score is significant at the $p<.05$ level.

**Difference between 1987 and 1989 score is significant at the $p<.01$ level.

Table 8

School Means for Teacher Attitude and School Climate

1987 and 1989

Level of implementation and school	Nonauthoritarian Attitudes		Respect for Students		Clarity of Rules		Fairness of Rules	
	1987	1989	1987	1989	1987	1989	1987	1989
Low (comparison)								
School 1	2.61	2.63	.95	.98	.75	.73	.48	.49
School 2	2.52	2.43	1.09	1.15	.82	.82	.73	.75
Medium								
School 3	2.48	2.74	1.07	1.08	.83	.82	.63	.59
School 4	2.38	2.58	1.01	1.04	.70	.75	.50	.50
School 5	2.48	2.34	1.01	.95	.78	.79	.53	.56
High								
School 6	2.62	2.75	1.11	1.15	.78	.75	.57	.57
School 7	2.41	2.75*	.88	1.07	.74	.85	.50	.66
School 8	2.39	2.46	1.08	1.03	.80	.78	.66	.64

Note. Bold entries indicate that the difference between the 1987 and 1989 climate scores exceeds twice the standard error of measurement.

* Difference between 1987 and 1989 score is significant at the $p<.05$ level.

Table 9

School Means for Measures of Student Experiences and Attitudes, 1987 and 1989

Level of implementation and school	Rewards		Positive Self-concept		Attachment to School	
	1987	1989	1987	1989	1987	1989
Low (comparison)						
School 1	.34	.33	.72	.73	.59	.59
School 2	.31	.38**	.77	.77	.74	.75
Medium						
School 3	.32	.34	.76	.74*	.72	.67**
School 4	.33	.39**	.71	.71	.63	.61
School 5	.30	.38**	.74	.70**	.63	.61
High						
School 6	.28	.35**	.73	.75*	.65	.66
School 7	.34	.33	.73	.74	.59	.68**
School 8	.36	.39	.77	.78	.75	.70**

* Difference between 1987 and 1989 score is significant at the $p<.05$ level.

**Difference between 1987 and 1989 score is significant at the $p<.01$ level.

Table 10

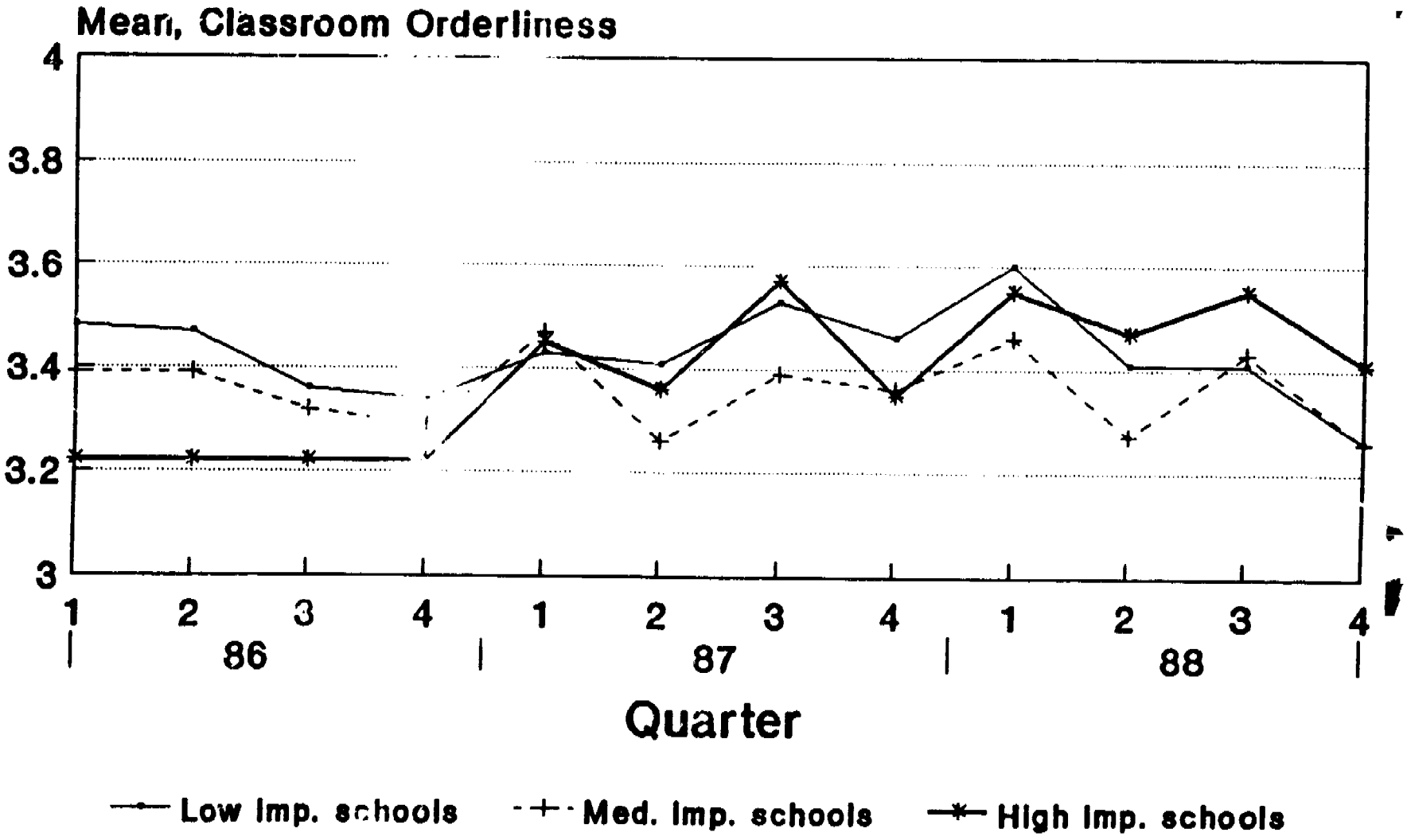
Summary of Effect Sizes from Effective School Battery Measures, by Level of Implementation

Outcome measure	Effect size for 1987-1989 change: All treatment	High implementation	Medium implementation	Low implementation (comparison)
Rebellious Behavior	.26	.09	.43	.19
Avoidance of Punishment	.05	.19	-.09	.06
Nonauthoritarian Attitudes	.21	.27	.15	-.05
Respect for Students	.18	.40	-.04	.25
Clarity of Rules	.28	.33	.22	-.25
Fairness of Rules	.24	.52	-.04	.17
Rewards	.08	.09	.18	.10
Positive Self-concept	-.01	.09	-.12	.00
Attachment to Schoo	-.05	.03	-.14	.03

Note. "Effect size" for climate measures (Respect for Students, Clarity of Rules, and Fairness of Rules) is the difference between the climate scores for 1989 and 1987 divided by the standard deviation for schools reported in the User's Manual (G. Gottfredson, 1985, page 57). For all other measures, the effect size is the difference between the two scores divided by the standard deviation for the school.

Figure 1

Student Reports of Classroom Order by Level of Implementation, 12 Quarters



First 4 Quarters are Baseline