In order to discover whether there were any inner-city elementary schools in Massachusetts which were successful in teaching children to read, a two-year study was undertaken. Results of the study, reported in this document, showed that there is no one pattern for success but that there is a process by which schools, either alone or in collaboration with other schools, can assess their strengths and weaknesses. This summary of the study outlines the problem, approaches, procedures, findings, and recommendations. (JM)
SUCCESS AND FAILURE

A Summary of Findings and Recommendations
for Improving Elementary Reading
in Massachusetts City Schools

A Study for the Massachusetts Advisory Council on Education

ALLAN B. ELLIS

Educational Research Corporation
85 Main Street
Watertown, Massachusetts 02172

February 1975
Educational Research Corporation is a non-profit research and development organization founded in 1938 "to further the cause of education by the conduct of experiments, studies, investigations, and research, and by the provision of research consultation to educational institutions." Composed of over forty fellows and members—educational researchers, statisticians, technologists, curriculum developers, psychologists, test and measurement experts—ERC's research projects cover the full range of educational interests and concerns. These fellows and members of the corporation serve as consultants on an as-needed basis to the full-time resident professional staff. Current clients located throughout the country include Federal and State Governments, local school districts, colleges and universities, and leading business organizations.

The Massachusetts Advisory Council on Education is an independent state agency created by special legislation (General Laws, Chapter 15, Section 1H) for the purpose of recommending policies designed to improve the performance of all public education systems in the Commonwealth. As such the Advisory Council provides support for studies which will recommend policies promoting and facilitating the coordination, effectiveness, and efficiency of these systems.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword</td>
<td>2</td>
</tr>
<tr>
<td>Summary</td>
<td>4</td>
</tr>
<tr>
<td>The Problem</td>
<td>4</td>
</tr>
<tr>
<td>The Usual Approach</td>
<td>4</td>
</tr>
<tr>
<td>An Alternative Approach</td>
<td>6</td>
</tr>
<tr>
<td>The Study</td>
<td>10</td>
</tr>
<tr>
<td>Procedures</td>
<td>11</td>
</tr>
<tr>
<td>The School Factors</td>
<td>12</td>
</tr>
<tr>
<td>Collective Judgments</td>
<td>15</td>
</tr>
<tr>
<td>The Findings</td>
<td>16</td>
</tr>
<tr>
<td>The Deviant Schools</td>
<td>21</td>
</tr>
<tr>
<td>Comparing the Schools</td>
<td>23</td>
</tr>
<tr>
<td>Contrasting the Schools</td>
<td>26</td>
</tr>
<tr>
<td>Next Steps</td>
<td>28</td>
</tr>
<tr>
<td>Recommendation for Action</td>
<td>28</td>
</tr>
<tr>
<td>Preconditions for Action</td>
<td>30</td>
</tr>
<tr>
<td>Mechanism for Action</td>
<td>30</td>
</tr>
<tr>
<td>Notes</td>
<td>31</td>
</tr>
<tr>
<td>References</td>
<td>32</td>
</tr>
</tbody>
</table>
FOREWORD

Why do some inner-city elementary schools succeed in the job of teaching children to read while most others fail? Despite what several national studies have suggested about the predominate influence of the home, are there not some school factors in the teaching of reading that are also important? If so, what are they? And, most importantly, can these factors be transported to other schools?

Over two years ago the Massachusetts Advisory Council on Education became interested in these questions and their possible answers. While a good deal of this interest was sparked by a study suggesting that there were identifiable school factors that did make a difference, a large part of this interest was brought about by the continued evidence that children in inner-city schools in Massachusetts were not learning to read or, at least, not learning to read well. The reading scores on the tests given by the Massachusetts Department of Education in 1971 clearly demonstrated the failure of most inner-city schools to effectively teach reading.

To find out if there were “successful” inner-city schools teaching children to read and, if there were, what could be learned from them that would be of value to other inner-city schools, the Advisory Council contracted with Educational Research Corporation. The hope was for a prescription for reading success in inner-city schools. The answer, after over two years of study, observation, interviewing and analysis, is, frankly, mixed.

It is mixed for the reason that while the study did, indeed, find “successful” schools and identity factors related to their success, they also found that there is no one pattern for success. To say this in another way . . . there is no single reason for certain schools being successful and other schools not. At this point we cannot say there is one prescription for teaching inner-city children to read well.

What we can say (and what the study recommends) is that there is a way . . . a process by which schools alone or in collaboration with other schools, can begin diagnosing their own strengths and weaknesses. From this analysis schools can then attempt to create and implement strategies for improvement. These strategies can cull from this summary and
its technical accompaniment suggestions as to what changes schools can introduce into their reading programs and plans.

Obviously, all the answers are not in this study. The Council hopes, however, that readers will find in this report ideas and suggestions that will be of practical here and now use to schools. As always, the Council appreciates the feedback of readers on the usefulness of this report.

Allan S. Hartman
Associate Director
Advisory Council on Education
SUMMARY

THE PROBLEM

Go into a city, find where the poor people live, visit one of the elementary schools their children attend, and the overwhelming likelihood is that you will be in a school that is failing to teach its students to read. Some people use common sense to explain this failure: surely no reasonable person can expect a school to do a good job of teaching children whose lives are surrounded by relentless poverty, overcrowdedness, malnutrition and disease, crime, welfare, a high cost of living, and too much noise; a school can accomplish very little in the face of such handicaps. Should someone be skeptical of this explanation—for instance, should he be uneasy about blaming children for what might better be seen as the school's failure—then he must contend with the surveys and technical studies of the past eight years from which some of the country's most respected researchers have drawn similar conclusions.

The Usual Approach

Foremost among such studies is that of the U.S. Office of Education, usually called the Coleman study after the principal author of the survey report, wherein the group of over sixty education specialists who conducted this nationwide project concluded that “schools bring little influence to bear on a child's achievement that is independent of his background and general social context.” (Coleman, et al., 1966, p. 325) And this is not the strongest statement on the matter. Indeed when Marshall Smith performed a subsequent, more exhaustive analysis of the very same survey data, he discovered certain “mechanical errors” in the Coleman Report which led him to conclude that to ascribe “little” influence to schools is to attribute too much influence to schools. (1972, p. 269) Then there is Christopher Jencks, another reanalyzer of the Coleman data who, in one place, states when schools with economically and racially similar students were compared, differences in school policies and
resources were rarely associated with pedagogically significant or statistically reliable differences in verbal achievement. (1972a, p. 70)

in another place, asserts that “neither school resources nor segregation has an appreciable effect on either test scores or educational attainment.” and further on, concludes that whatever differences we may notice among schools “seem to have very little effect on any measurable attribute of those who attend them” (1972b, pp. 7-8) We need go no further to make the point that there exists a body of research literature which asserts that once certain non-school factors in the lives of children are taken into account, there can be found little relationship between what schools do and what students achieve.

But there is another side to the story. Based on a different brand of common sense, other people argue that Coleman and the rest are wrong: the consistent failure of poor city children to learn and to achieve is not the result of what the children are, but of what the schools fail to do for them. Consider, for example, a conclusion reached by the Task Force on Urban Education:

Many systems' unconscious biases and static expectations have limited their capacity to teach children who enter the schools without certain attributes held by previous constituencies of the system. Such attributes relate to being oriented to middle class values and expectations, being reading-ready, and having the structural orientation that facilitates shifting from subject matter to subject matter as dictated by time blocks rather than by interest and substance. Because of the widespread use of systems' equating a student’s capacity to meet their expectations with his possession of such middle class attributes, the concept of the self-fulfilling prophecy has all too often been demonstrated. (1970, p. 34)

Thus the prophecy of failure will be fulfilled and after six years of schooling the relative ability of the typical urban student to do school work will have diminished not because the raw material is deficient but as Kenneth Clark has said, because children who are treated as if they are uneducable invariably become uneducable. (1965, p. 128)
Nor is common sense the only basis for the belief that schools can affect the achievements of its students. Looking again to the educational research community, we find no lack of surveys and studies drawing conclusions directly opposite to those of Coleman, et al. For instance when James Guthrie reviewed nineteen projects which analyzed the relationship between school components and pupil performance, he became sufficiently impressed with the amount and consistency of evidence supporting the effectiveness of school services in influencing the academic performance of pupils to conclude, "there can be little doubt that schools do make a difference." (1970, p. 46) Just a few of the surveys Guthrie studied are those of Mollenkopf and Melville (1956) involving 17,000 ninth and twelfth grade students throughout the United States, Goodman (1959) which looked at 70,000 seventh and eleventh graders in 102 school districts in New York, Burkhead (1967) which contained 109,000 students from 241 schools in Chicago and Atlanta, and Guthrie himself (1970) which focused on 5,234 sixth graders in Michigan. The most recent survey of the relationship between schools and pupil performance, which involved "complete data on all fifty states," supports Guthrie's conclusion and led the authors to assert,

... we believe one implication drawn from the Coleman Report—that variations in education inputs make little difference for test performance and equality of opportunity—is scientifically and socially unsound. (Walberg and Rasher, 1974, p. 9)

And so it goes.

An Alternative Approach

Now if competing common sense interpretations of the matter, each bolstered by conflicting research surveys and opposing schools of expert opinion, were all we had to go on, not only would prospects for resolving the issue be remote, but those practical decisions about how to better teach students, which must be made every day in the real world of schools, would have the strange character of depending not so much on research, itself, as on which researchers the deci-
sion-makers are persuaded to believe. But the untenable position in which contradicting “findings” place schools and the people who run them, and the need to establish a firmer base on which to make educational decisions, have combined to make schoolmen receptive to an alternate approach which begins with the notion that, notwithstanding incontrovertable evidence of the general failure of inner-city schools, some urban schools, somewhere, are succeeding to teach the very same type of children that the other schools are failing. Find even one such school, the notion says, go look first hand at what they do there, and see if you can tell what makes it so different from the usual city school.

The most widely known study of this type is the one George Weber performed during the 1970-71 school year which he describes in a monograph entitled, “Inner-City Children Can be Taught to Read: Four Successful Schools.” Weber knew that reading achievement in almost all inner-city schools “is both relatively and absolutely low,” and he knew “most laymen and most school people believe that such low achievement is all that can be expected,” but as well he believed there were inner-city schools “in which reading achievement was not relatively low, in which it was, indeed, about the national average or better.” (1971, p. 1) Thus his purpose became to test this belief, and should it prove to be correct, to “show that inner-city children can be taught reading well,” while at the same time—hopefully—to “discover some common factors in the success of the good programs.” (p. 2)

The Weber study located four elementary schools, each of which met the criteria of a) being a non-selective public school in the central part of a large city and attended by very poor children; and b) having a median pupil achievement in reading which matches the national norm, with a low percentage of gross failures. (p. 5) Weber visited each school, spending two or three days observing teachers and classrooms, and interviewing the administration and staff. On the basis of these visits and his subsequent analysis, Weber concluded there are eight school factors that “seem to account for the success of the four schools:” strong leadership, high expectations, good atmosphere, strong emphasis on reading, additional reading personnel, use of phonics, individualization, and careful
evaluation of pupil progress. (p. 30) Underneath the specifics of Weber's findings, of course, is the more general realization that the very fact of locating and documenting the existence of these successful schools "shows the failure in beginning reading typical of inner-city schools is the fault not of the children or their background—but of the schools." (p. 30)

The logic of Weber's approach is compelling. That is, putting aside his findings and conclusions, his method of addressing the problem is appealing to common sense and responsive to the need to be practical about how we fashion solutions. For instance, in response to massive, nationwide surveys involving hundreds of thousands of students, where usually the variables are proxies—indirect measures or stand-in approximations—and the data are collected through questionnaires and other second-hand means, Weber proposes we go to a few schools and look at what goes on, thus replacing the survey method with the case method. In response to the usual regression analyses, data transformations, tests of significance and other statistical procedures applied to number-filled matrices, Weber proposes we examine people's behavior directly, thus replacing the statistical approach with the clinical approach. In response to the measures amenable to quantification usually found in surveys—pupil/teacher ratio, size of school, number of books in the library, teacher salaries, demography of students, size of science laboratory facilities—Weber proposes such things as leadership, atmosphere, individualization, strong emphasis on reading, and the like, thus replacing quantity with quality.

To be fair the educational researchers who do surveys are well aware of the strengths in the general approach which characterizes the Weber study. In his reanalysis of the Coleman data, for example, Smith questions certain of the variables used when he says

> It is possible, however, that a survey is an ineffective way of measuring differences among schools in facilities and curriculum resources. It may be that data about the quality of science facilities are more important than data about their quantity. (1972, p. 282)

Similarly, in supporting or strengthening the Coleman, et al. findings, Jencks still raises questions of the accuracy of data used as estimates of school policies and resources, and con-
eludes "data on different school or student characteristics might well yield different results." (1972a, p. 73) Then there is Henry Dyer who counsels us to seriously entertain the hypotheses "that measures used in [Coleman, et al.] to describe the schools and their functioning are simply too crude to yield any useful amount of meaningful variance." (1972, p. 385)

At the same time, of course, any trained educational researcher is able to spot serious weaknesses and raise important questions regarding the way Weber chose to carry out his study. The small size of the sample, his use of just one observer in the schools, questions about his observation procedures, the suitability of his output measure, and the bases upon which judgments were rendered and commonalities among schools identified all come easily to mind as concerns or potential problems with Weber's study.

But the most severe problem with the study is the absence of any comparison schools on which to double check the conclusions and generalizations offered. Studies that confine themselves to looking at successful schools to see what makes them succeed run the risk that the school factors they identify as being related to excellence may very well exist either to varying degrees or with equal potency in failing schools. That is, it may not be strictly a matter of the presence or absence of factors that makes the difference but either one of magnitude or of pattern. As a consequence of his approach, however, the inference Weber encourages, intentionally or not, is that good schools possess all eight characteristics he identifies, and poor schools have none. Yet certainly there is no clear basis for such an inference, since we are given no information—Weber's study produced none—on how poor schools stack up on his factors. If anything, experience should be enough to suggest that things are not likely to be so simple. So despite the general appeal of the approach used by Weber, we are not so well off because of the study as many readers of his report have supposed.
THE STUDY

One of the first groups to recognize that there is considerable strength in the general approach Weber adopted, but certain weaknesses in his specific implementation is the Massachusetts Advisory Council on Education. MACE realized, as well, that the Weber study, because of its intuitive veracity and its sensitivity to the critical issues facing schools, was likely to persuade and influence decision-makers in schools, despite the need first to resolve various questions left unanswered by the study. Convinced of the need to conduct a well-formulated study within the Commonwealth, employing the general approach of looking at successful inner-city schools to see what makes them work, MACE called upon Educational Research Corporation to help construct a study of this type that would have several critical features, among them:

- six to ten successful, inner-city elementary schools in Massachusetts were to be examined to allow reasonable breadth and generalization;
- an equal number of failing schools—call them contrast schools—would be included in the analysis to permit the identification of factors unique to the successful schools;
- a practical orientation such that, in light of any findings, there would be concern for what schools might begin to do to improve;
- the involvement of experts from various relevant fields: reading, measurement, administration, individualized instruction, and such and a concern for systemizing the process of achieving consensus of judgments among experts;
- the judicious combination of clinical and statistical methods;
- the involvement, as advisors, of policy makers within the state and reading practitioners from the major cities.
Procedures

Following this plan, the MACE Study of Reading—conducted by Educational Research Corporation and begun in January 1973—consisted overall in the visiting, observing, and characterizing of twenty public, non-selective, inner-city elementary schools in the Commonwealth. These schools were selected on the basis of standardized tests of reading achievement administered to sixth grade students, such that the average performance of the students in ten of the schools placed them at or above grade level compared to national norms. In the other ten schools, students had reading scores which, on the average, were 1.3 grade equivalents below national norms. All twenty schools had high percentages of students from poor families. Each of the ten “successful” schools was paired with a “contrast” school that matched it on the poverty measures as well as racial composition and proportion of bilingual students. Situated in nine cities throughout Massachusetts, with school populations varying from 10,000 to 100,000 students, some housed in old buildings with self-contained classrooms, others located in newer buildings with open spaces, the twenty schools studied represent a reasonable cross-section of city schools within the Commonwealth.

A group of twenty-five professionals—specialists in reading, teaching, observation procedures, language arts, and such educational research areas as measurement, instrument development, statistics, attitude survey, and data analysis—assisted Educational Research Corporation in preparing for and conducting the visits to the schools. Various teams of three or four specialists drawn from this group made a total of five visits to each of the twenty schools. So as not to bias the observations, no member of the visiting team knew whether the school had been classified successful or contrast. In each school the visiting team did the following things:

- systematically observed ten reading periods;
- conducted structured interviews with the principal, five reading teachers, the reading specialist, the school librarian, and the central office staff responsible for reading in the district;
collected background information from the principal, the school reading specialist, and all the teachers in the school:

- collected three days worth of time logs showing classroom activities from each teacher of first, second, and third grades:
- administered a test of knowledge of children's literature to all teachers of fourth, fifth, and sixth grades:
- collected school-wide demographic data:
- administered a forty-five item attitude toward reading inventory to a random sample of sixth grade students;
- administered a home background questionnaire to another random sample of sixth grade students.

The School Factors

Each datum collected out of all these activities was keyed to one of the eleven operationally defined, qualitative school factors whose relationship to student performance the project was designed to study. (Note 1) The factors and their definitions follow:

Leadership in a school requires having one or more persons who provide direction to the activities of the staff in the teaching of reading. This person or persons is generally in a supervisory role and displays characteristics such as inspiration, empathy, flexibility, and the like. When there is clear leadership in reading, the staff will consistently agree on who provides the direction, whether it be the principal, a school reading specialist or someone from the central office.

Coordination of the various activities in the teaching of reading to ensure that there is continuity across and within grade levels can be considered a function of leadership, but it was deliberately separated for special attention. The study staff deemed it important to verify that a student's work at any grade was related to the experiences of earlier grades and that the several supplementary reading services, remedial, Title I, or learning disability, which take place outside of the classroom are still related to the work in the classroom. Coordination can be achieved simply by standardization of all
activities, but a school with varied learning activities must see that the activities are well orchestrated.

Additional reading personnel include those other than classroom teachers who actively work in the reading program, the most common example being the school reading specialist. The roles of reading specialists may differ, for they carry a remedial caseload, or serve as the coordinator of the school reading program, or both. Other possibly important personnel may include the Title I staff and the learning disability specialists, as well as paid or volunteer aides, intern teachers, older students, and so on. In some cases the school librarian may perform as an additional resource person for reading.

Atmosphere of a school can be described by many dimensions, and Weber listed several adjectives that he found to be applicable: orderly; purposeful; quiet; and relaxed. We decided to use these objectives to create scales, e.g. from orderly to disorderly, quiet to noisy, and so on. It was agreed that care must be taken to distinguish between different types of noise since many constructive activities are just naturally accompanied by noise.

While Individualization means quite different things to different people, there was general agreement that individualization should be responsive to student differences in learning and that there should be a diagnostic program to identify such differences. The individualization process itself might provide for students to work on individually prescribed materials, or it might allow students to proceed at different rates, or a combination of the two. Given the diversity of individualization techniques used, we determined to identify whichever ones happened to exist in the study schools.

Evaluation of pupil progress requires collecting performance data regularly, but the means for collecting are many: teacher constructed tests; curriculum tests; criterion-referenced tests; standardized tests; and so on. For the evaluations to be an effective tool in developing instruction strategies for a student, there must be evidence that the reading progress records follow that student as he or she moves from grade to grade and from teacher to teacher.

High Expectation. Teachers differ on the performance levels which they think their students are capable of attain-
ing, and, as a consequence, they can also differ on the standards that they set for their students. Expectations can differ among teachers according to how they feel about their particular students, but they can differ as well according to how they feel about students in general.

Strong Emphasis On Reading. There are many ways in which a school can place strong emphasis upon reading in contrast to other school activities. Among them are time devoted to reading instruction and the amount of reading materials made available. Moreover, there are more subtle indicators of the commitment to reading; as, for example, whether or not reading is taught at the beginning of the day when the children are most alert.

Use of Phonics. The importance of developing decoding skills in the early years has become so widely recognized that most schools are likely to use phonics in their program. Therefore, the important concern is not the presence of phonics but rather how central to the reading instruction the use of phonics actually is. The study determined to focus upon the reading materials used by the schools to find out the extent to which phonics was the major point of focus.

Staff Training and Experience. Teachers and others in schools have different backgrounds in terms of their own education and of their experience in education. Important dimensions are the amount of formal education, including particularly the number of reading courses taken at a college or university or as part of an in-service program, and the number of years spent in the education profession.

Quality of Teaching. The study determined that the quality of teaching was not a critical factor in the successful schools, where he found competent but not outstanding teachers. Given the logical possibility that poor teaching probably results in poor performance, the study staff decided to check on the extent to which the contrast schools may be less successful because of poor teaching. Appraising quality of teaching is quite presumptive and requires some limits to keep the process as objective as possible. We decided upon observing the management of the classrooms, the interactions between teacher and pupils and, the degree to which students engaged in learning activities when they were on their own in work sessions at their desks.
Collective Judgments

Utilizing clinical and statistical procedures in combination, the study employed an analysis technique which allowed the various specialists to generate collective judgments about all the schools. They rendered these collective judgments factor by factor in a thorough and systematic way: after extensive analysis and discussion of all the data, the group developed a consensus rating of each school on each factor, indicating whether in its judgment the data relative to that factor suggested that the school was successful or unsuccessful. These judgments were made on a scale where a rating of “1” meant on this factor “appeared definitely to be a contrast school,” a “5” meant “appeared definitely to be a successful school,” and a “3” was the point of equal probability, given the data. This procedure—which resulted in eleven factor ratings for each school—was possible because none of the specialists was told whether the standardized test scores of a school placed it in the category of successful or contrast school.
THE FINDINGS

The year long development of plans, strategies, and instruments, the extensive visits to schools by teams of experienced professionals, the careful analysis of all the data, the vigorous discussions to achieve consensus ratings, and, finally, the application of just about all the major procedures for comparing and contrasting groups led up to the posing of the single, basic question of the study: are there discernible differences between the group of successful schools and the group of contrast schools on any or all of the eleven factors, so as to explain the differences between the two groups on achievement test scores? In our attempt to answer this question we discovered not two but four groups of schools and this discovery holds several important implications for how city schools can view themselves and their strategies for change. Inspection of the factor ratings, school by school, revealed that some of the successful schools look more like contrast schools and some contrast schools look more like successful ones.

Thus the study contained four types of schools, as follow:

1. True positive schools generally rated high on the eleven factors, and where student test scores are high;
2. False positive schools generally rated high on the eleven factors, and where student test scores are low;
3. False negative schools generally rated low on the eleven factors, and where student test scores are high;
4. True negative schools generally rated low on the eleven factors, and where student test scores are low.

The following picture will help to visualize these four types of schools. The figures in the lower right corner of each cell represent the number of schools from the sample that are of each type. (Note 2) Looking first at each cell separately reveals some interesting facts and raises some intriguing questions about the four types of schools.

The Successful School (True Positive: Cell 1). The schools in this cell come closest of all those in the sample to being of the type Weber and others refer to when they write about the successful inner-city school. Yet, when we examine the factor ratings these schools received, the picture is more varied. Uniformly the true positive schools place a
## Test Scores

<table>
<thead>
<tr>
<th>Factor Ratings</th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>CELL 1</td>
<td>CELL 2</td>
</tr>
<tr>
<td>HIGH</td>
<td>TRUE POSITIVE (Successful)</td>
<td>FALSE POSITIVE (Transition)</td>
</tr>
<tr>
<td>N = 4</td>
<td></td>
<td>N = 5</td>
</tr>
<tr>
<td>LOW</td>
<td>CELL 3</td>
<td>CELL 4</td>
</tr>
<tr>
<td>FALSE NEGATIVE (Ersatz)</td>
<td>TRUE NEGATIVE (Unsuccessful)</td>
<td></td>
</tr>
<tr>
<td>N = 6</td>
<td></td>
<td>N = 3</td>
</tr>
</tbody>
</table>

*strong emphasis on reading,* but this is the only factor they have totally in common. Viewed as a group, however, they show the strong tendency to be high on ratings for atmosphere, staff training and experience, and quality of teaching—this last factor having been found by Weber to be “not essential” to a school’s success (p. 30), and while true positive schools are uniformly low on none of the factors, they have the tendency to be weak on coordination and individualization. The overall character of these schools is quite positive, to be sure, but even within the areas of common strengths, variation abounds. Anecdotal reports from the four visiting teams paint a similarly varied picture. With respect to reading classes, for instance, what the visitors observed ranges from “rooms colorfully decorated with student work; pleasant, friendly, yet structured and controlled atmosphere;” and “the presence of special personnel to observe classroom work of students who are candidates for work in the Learning Center;” to “generally barren classrooms;” “detached, preoccupied teachers;” and “emphasis on recall with no especially probing questions being posed.”

Thus notwithstanding the study’s very clear sense of these schools as truly successful, as models for what schools should
be like, they pose several problems, two of which are especially telling. First, as we have seen, none of these schools is strong on all the school factors, thereby raising serious doubts about the inference drawn by the typical reader of Weber that all factors are directly involved in a school's success. Second, moreover, successful schools are not all strong on the same factors. That is, while there are certain common tendencies toward excellence that can be identified, each school has a different overall pattern of strengths and weaknesses on the factors, suggesting—as we guessed from the start—things are not quite so simple as prior studies of this kind have allowed us to believe. Success by formula or even by mere imitation is not likely to be possible.

The Transition School (False Positive: Cell 2). The schools in this cell uniformly place a strong emphasis on reading and display a high quality of teaching. As a group these schools show some tendency to be high on staff training and experience, leadership, atmosphere, and the use of additional reading personnel. They are uniformly low on no factors, but have a strong tendency to be low on individualization and expectation, and some tendency to be low on coordination. Characterizing the false positive schools is difficult since, on the face of it, they are doing a number of desirable things as indicated by their ratings on the school factors, and yet the performance of their students does not reflect the force of these strengths.

The hypothesis that it is a matter of time before these strengths will show their effects is appealing and would be consistent with Weber's finding that the inner-city schools he studied required, in some cases, as much as nine years to achieve success. Thus we might view the false positive school as a school in transition. But the absence of a common pattern among the school factors of these schools—an absence which persists throughout the cells—and our general uncertainty about the direction of any change these schools may seem to be undergoing, makes caution the critical ingredient of any guesses we may be persuaded to make. Only had this study used a longitudinal design—time and money did not permit it—where growth or decline could have been better charted could we have begun to observe with any clarity the direction and magnitude of change. Still there is sufficient
evidence to warrant carrying the notion of transition to subsequent sections of this summary where the question will be addressed again.

The Ersatz School (False Negative: Cell 3). The schools in this cell are uniformly low on coordination, evaluation, and the use of additional reading personnel. As a group they show the tendency to exhibit poor leadership, to lack individualization, and to hold low expectations for their students. Over half the group is rated very low on atmosphere, and there is no school factor on which they are high or even tend toward being high. The dominant themes that persist throughout these six schools are those of discipline and drill. Excerpts from the observation reports of the six teams that visited these schools present a disappointingly consistent picture:

- The prevailing attitude toward the students is one of a we-they relationship. Teachers seem to view the children almost as the ‘enemy.’
- Teachers are continually distracted by keeping classes quiet, demanding students to pay attention and get to work.
- There is a very strong orientation to drill and practice. Much of class time is spent having the entire group go over worksheets, or generally drilling students on basic skills.
- There seems to be an overemphasis on phonics and, there were several instances of the teacher being unduly critical of dialect differences.
- There is a strong commitment to discipline as “the only thing the students understand.”

For the most part, these schools are drill schools where, seemingly, little attention is paid to how reading skills are used either to obtain and process information or to gain pleasure.

The very best judgment, systematically rendered, of a group of education specialists awards these schools the extremely strong likelihood of failing, and their students end up achieving relatively high test scores. By all appearances the false negative school is a contradiction. But because one does
not easily discount such expert judgments, the persuasion is to ascribe the difficulty to a faulty placement of these schools among the successful schools. Yet there are six such 'faulty placements' and, indeed, their students do have sufficiently high test scores to warrant their assignment to the successful pool. But perhaps they were wrongly placed in another sense. The strong orientation among these schools to drill, coupled with the general absence of classroom time devoted to the use of basic reading skills suggests that the students' ability to perform well on skill related standardized tests may not be indicative of their ability to utilize reading skills in all the various ways normally associated with the designation 'reader.' Thus, perhaps standardized test scores constitute an insufficient criterion with which to judge a school to be successful.

Now the critical question at this point is not at all concerned with whether or not our guess about these Ersatz schools is found to be persuasive; the data do not go far enough. Instead we focus on what we would have to do to tell with some certainty whether these schools warrant assignment as successful. The obvious solution is to use a more appropriate measure of success. The presence of false negative schools in the sample suggests that just as Coleman and the other surveyors are wrong to use measures of quantity to characterize inputs to education, so this study is wrong to rely solely on quantitative measures of school outputs. There is nothing new, of course, about the notion of satisfactory performance on standardized tests of reading skills being a necessary but not sufficient condition for reading excellence. Indeed once we go so far as to recognize the need to replace quantity with quality in our measures of school input, it becomes nearly ironic that we should miss the parallel of seeking measures of the quality of student performance. But this conclusion is not offered to explain the presence of these schools; advocates of the recent and spreading movement toward "back to the basics" alternative schools in such places as Pasadena, Palo Alto, Cupertino, California, Charlotte, North Carolina, Lakewood, Colorado, and Prince George's County, Maryland might provide an alternate explanation. The intent is to view measures of output quality—desirable in any case—as a means to resolve the apparent contradiction posed by the Ersatz schools.
The Unsuccessful School (True Negative. Cell 4). These schools in this cell are uniformly low on evaluation, coordination, individualization and the amount of emphasis placed on reading. In addition they show a strong tendency to hold low expectations for their students, to make poor use of additional reading personnel, to manifest poor leadership, and to possess among the staff a minimum amount of training and experience. Naturally there are pockets of quality among these schools, but the overwhelming tendency is one of inadequacy. Indeed there is no single factor on which they are uniformly high or even on which they display some tendency to be high. Mediocrity is the highest level of rating these schools attain.

The Deviant Schools

These cell descriptions offer a beginning sense of the four classes of schools this study has identified. This beginning sense will be amplified and enhanced further on where we compare the cells to see what they have in common and what is unique among them. Then, by contrasting certain cells with each other, we can return to a variant of the original question of how schools succeed—maintaining the spirit of the concern that motivated MICE to establish this study, but adding an important constraint imposed by the existence of four types of schools, which will allow us to touch briefly upon the process of becoming a successful school. Before turning to these tasks, however, there remains one aspect to explore regarding the way we have come to categorize the schools that participated in this study.

The four categories of schools are not mere artifacts of this study. Of course the study’s capacity to perceive these four categories depends upon its having begun with a set of factors whose relevance we wished to explore. Had we proceeded as have others, attempting to derive the factors by contrasting good and bad schools, we would not have noticed the applicability of the four-cell approach. But since the study posited a set of factors presumably related to student achievement, it became possible to classify the schools in terms of both factor ratings and achievement scores, this
joint function implying the four combinations represented by the cells. Put another way, since the hypothesis to be explored was that somehow or other a school's standing on the eleven factors is related to its success or failure, there is engendered by this hypothesis a general expectation—the veracity of which the study was to test—that high factor ratings are accompanied by success and low ratings spell failure. Thus what might otherwise be considered an aberrant school, in that it does not look like the others in its class of succeeding or failing schools, becomes for us a school that does not live up to that expectation implied by the study's hypothesis. When it becomes possible, thereby, to separate these deviant schools into those that exceed and those that fall below expectations—Ersatz and Transition schools respectively—there is formed the basis for a four cell category scheme.

Now this scheme is useful for characterizing the study schools, but its significance rests as well upon two inferences which we have drawn earlier concerning the deviant schools. From the schools falling below expectation (the false positives) we have drawn, with caution, the inference that they are transition schools. Up to now we have spoken of observing and rating certain school characteristics and processes to discover what good schools are like. But given there are schools in transition, perhaps we should focus future observations upon the strategies and procedures for change being utilized, whether implicitly or explicitly, by these transition schools. That is, just as we can observe and rate school characteristics and processes, so perhaps we may look to these transition schools to discern and classify the strategies which direct or account for their movement toward or away from excellence. From the other set of deviant schools, those exceeding expectation (the Ersatz schools) we draw, with some confidence, the second inference that measures of school output should reflect qualitative as well as quantitative concerns. But naturally before such measures can be specified, schools must make explicit the particular educational objectives the attainment of which is to be measured. Only when the relationship between these objectives and the devices used to measure student progress has been established can there be any certainty concerning the relevance of such measures. Yet it is not now common practice for schools to enunciate
their objectives unambiguously and to administer tests specifically designed to depict student growth toward these objectives, although the reader familiar with criterion-referenced testing will know this is not a new idea.

Comparing the Schools

The chart on the following page summarizes what has been said so far about the four types of schools the study identified. On this chart the columns represent the four cells as they have been defined earlier, and the rows represent the continuum of school factor ratings. Inspection of the chart reveals some interesting things about the study schools. Notice the movement on the chart, overall, from high to low ratings on the school factors, with the positive schools having very few low ratings and the negative schools having no high ratings. But since all four types of schools are rated low on coordination and individualization, these factors can be dropped from consideration because the central concern here is with contrasting the schools to see how they differ. Thus the positive schools, as a group, show only one factor on which they are low—expectation—a fact which will become important as we look at similarities and differences between the two types of positive schools.

Overall the positive schools uniformly place a strong emphasis on reading and provide their students a high quality of teaching. Whereas the Successful schools achieve high ratings on atmosphere and staff training and experience, the Transition schools are not there yet, although they show a tendency to be high on these same factors. The Transition school rates somewhat higher on leadership and the use of additional reading personnel, but the interesting fact about these schools is they are the only ones showing a strong tendency in either direction, lending some credence, thereby, to the view that they are transition schools and that it is a positive direction in which they are heading. Still the most striking feature of these schools and the area of greatest difference from Successful schools is their extremely low rating on expectation. This low rating is surprising since one would expect a school especially a transition school with such
<table>
<thead>
<tr>
<th>Level of Ratings</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True (Cell 1)</td>
<td>False (Cell 2)</td>
</tr>
<tr>
<td><strong>Successful</strong></td>
<td>Emphasis on Reading</td>
<td>Emphasis on Reading</td>
</tr>
<tr>
<td></td>
<td>Quality of Teaching</td>
<td>Quality of Teaching</td>
</tr>
<tr>
<td><strong>TENDING HIGH</strong></td>
<td>Atmosphere</td>
<td>Training &amp; Experience</td>
</tr>
<tr>
<td></td>
<td>Training &amp; Experience</td>
<td>Leadership</td>
</tr>
<tr>
<td><strong>MEDIUM</strong></td>
<td>Leadership</td>
<td>Training &amp; Experience</td>
</tr>
<tr>
<td></td>
<td>Add'1 Rdg. Personnel</td>
<td>Evaluation</td>
</tr>
<tr>
<td></td>
<td>Evaluation</td>
<td>Phonics-Supplemental</td>
</tr>
<tr>
<td><strong>TENDING LOW</strong></td>
<td>Coordination</td>
<td>Coordination</td>
</tr>
<tr>
<td></td>
<td>Individualization</td>
<td>Individualization</td>
</tr>
<tr>
<td><strong>LOW</strong></td>
<td>Coordination</td>
<td>Coordination</td>
</tr>
<tr>
<td></td>
<td>Individualization</td>
<td>Individualization</td>
</tr>
<tr>
<td></td>
<td>Expectation</td>
<td>Expectation</td>
</tr>
<tr>
<td></td>
<td>Leadership</td>
<td>Leadership</td>
</tr>
<tr>
<td></td>
<td>Evaluation</td>
<td>Evaluation</td>
</tr>
<tr>
<td></td>
<td>Training &amp; Reading</td>
<td>Training &amp; Reading</td>
</tr>
</tbody>
</table>
excellent features as strong emphasis on reading, high quality of teaching, good leadership, appropriate use of additional reading personnel, and the like to be doing all these things with the expectation that they will make a difference in the achievement of students. Yet on the expectation factor, Transition schools look much more like the negative schools than like their positive partners.

To say the low expectation they hold for their students fully accounts for the failure of students in these Transition schools to achieve—and thus explain the difference between the true and the false positive schools—is probably saying too much, although such an argument would not be without persuasion. It seems more prudent to suggest, however, that the failure of the Transition school students is due mostly to the lag between improvement of a school's instructional delivery system and the effect such improvement has on the performance of students. Still it is difficult to imagine these schools ever becoming Successful schools without a marked change in their attitude about what students can accomplish, and the setting of correspondingly high standards for their students to achieve.

Turning now to the negative schools, the complete absence of factor ratings above mediocrity and their propensity to be so low on so many factors certainly distinguish this group from the positive schools. More specifically, again removing coordination and individualization from consideration, these schools consistently exhibit low expectations for their students, poor leadership, the lack of inappropriate use of additional reading personnel, and inadequate evaluation of students. The Unsuccessful schools are noticeably worse on emphasis on reading and staff training and experience and somewhat better on atmosphere. Otherwise the negative schools match each other more closely than do the positive schools. There seems little doubt about the Unsuccessful schools being true failures, although the Ersatz schools remain difficult to fathom—remember their students do well on reading tests—although the evidence suggests the original characterizations of these schools is on the right track: more and more the Ersatz schools give the impression of having been Unsuccessful schools that attempted to improve student achievement through drill and practice alone. Thus, it seems,
they emphasize reading by emphasizing little more than the reading skills measured by standardized tests.

Contrasting the Schools

Now if we assume the true positive schools are the best schools—forgetting for the moment, for example, that there is room undoubtedly for improvement among them as evidenced by their extremely low ratings on coordination and individualization (Note 3) then we can look at their "medium" ratings and get an interesting hint about the force exerted by the school factors examined in this study. On the factors leadership, additional reading personnel, evaluation, and expectation the Successful schools are rated medium, meaning judgments about these schools, were they based solely on performance on these factors, would be inconclusive. Apparently, then, mediocre performance in these areas is not sufficient to hamper the true positive schools' ability to succeed with their students. Yet, again discarding coordination and individualization, these same four factors account for all the low ratings of Ersatz schools and all but two low ratings for Unsuccessful schools. The suggestion here is that perhaps the factors, or certain ones of them, exert a threefold effect: their presence contributing toward excellence, their meagre presence having neutral effect, and their negative presence contributing to failure. That is, most if not all the factors have negative sides to them, such that a low rating does not mean the absence of behavior so much as its perverted presence. Thus low ratings on leadership might mean regimentation, low ratings on expectation might mean despair, and low ratings on evaluation might mean judgment and punishment.

Though conjectural at this point—although the notes and anecdotes from the visiting teams support this notion—this idea is consistent with the fact that the factor ratings of the Ersatz (negative) schools stand pretty much in the same relation to each other as do those of the Successful (positive) schools—the difference being one of absolute magnitude of the ratings. Therefore while we cannot look at the true
positive schools alone and determine that these "medium" factors are important to success, we may become impressed by the difference between their ratings and those of the negative schools on the same factors. The difference between the positive and negative schools on Leadership, for example, is almost as great as their difference on Emphasis on Reading and is greater than their difference on Quality of Teaching.
The MACE study of reading entailed the formulation and application of a systematic procedure for the characterizing and evaluating of an elementary school's instructional delivery system, specifically with respect to the teaching of reading. The findings of this study indicate *Emphasis on Reading, Quality of Teaching, Atmosphere, and Staff Training and Experience*, by virtue of their high position among Positive schools, to be good candidates to be the determining factors of success. Joining these factors by virtue of relative positions among Positive and Negative schools, are the additional factors, *Leadership, Expectation, Use of Additional Reading Personnel*, and *Evaluation*. But the study found, as well, that unlike the presumption of Weber and others, there is discernible no single pattern of school factors that determines excellence. Thus even though the factors used in this study appear reasonable and relevant ones with which to describe a school's performance, and along which a school can attempt to improve, there are no grounds on which to point to any particular one of the Successful schools as a model. Thus instead of imitating a model of excellence, we recommend that schools focus on the *process* by which they can achieve excellence, each school thereby establishing its own brand or pattern of factors.

**Recommendation for Action**

The underlying premise of the course of action we recommend to schools is that they should

1. look at themselves the way this study looked at them.
2. diagnose their strengths and weaknesses on the study factors.
3. attempt to create and implement strategies for improvement.

For the first step of the process, which involves observing school practices and interviewing school personnel, rendering consensus ratings on each of the school factors, analyzing student performance, and thereby establishing the school's category, we do not recommend the use of outside specialists.
While reading specialists, researchers, interviewers, and other such professionals enhance the depth and perception of observation, the school and its staff become cast in the passive role of being scrutinized, precisely at the time it would be best to involve them and make them accessories in the process of improvement. Thus it seems best for the staff to be involved directly in the collection of the data and in their analysis.

Yet much of the clinical data require an objectivity of perception that might be unrealistic to expect from staff. Those data are better collected by peers from neighboring schools in a collaborative manner, this collaboration falling between, for example, that of informal district meetings of teachers and that of the formal accreditation visits. Such collaboration, entailing several schools helping to observe one and other, would involve the collective staffs in a form of introspection undoubtedly good for its own sake. Moreover, during the rating of school factors by consensus, this collaboration would encourage a systematic yet broad conversation among school people on the desirability and relative effectiveness of various observed practices.

Upon completion of the first phase of the process each of the collaborating schools will know where it stands—in which cell it falls—with respect to its profile on the school factors and on the measures of student performance. Then the task becomes to select and implement strategies for improvement in the school factors,—is it best to minimize weaknesses, maximize strengths, or attempt both is a question whose answers will vary—and here the collaboration among schools is most critical since it is the combined best judgment of a collection of school people that is most likely to generate relevant and workable programs for systematic change over the long haul. Besides as the schools, through this process of self identification and improvement, become more like laboratories for the study of change, it is imperative that they monitor each other over time, paying particular attention to those schools identified as being in transition to calibrate and revamp strategies. Thus there is introduced into the process, a formative aspect with which to make the successive approximation more systematic.
Preconditions for Action

Now three steps must be taken before schools can profitably engage in the process described in this study.

1. The instruments of this study need to be revised to reflect greater sensitivity to there being four types of schools, and especially to the observation of strategies being utilized generally, but in particular in the transition schools.

2. Each school must determine for itself which outcome measures are relevant vehicles to assess the degree of its success. Probably this will entail stating explicitly and in measurable terms the objectives of reading instruction in that school. For some objectives, standardized tests of reading may be found to be appropriate, while for other objectives existing but non-standardized instruments—for instance those of the National Assessment Project, already in use in the Commonwealth through the State Department of Education—may be most appropriate. For still other objectives, unobtrusive measures of reading behavior may be desirable. In any case, the act of stating objectives and determining relevant measures of student attainment of these objectives is clearly an essential element in the process.

3. Arrangements need to be made for training school people in several facets of the process, among them systematic observation, interviewing, sorting and analyzing the data, and especially in the consensus process where judgments on the various school factors are rendered, scrutinized, discussed, and a final rating determined.

Mechanism for Action

The Massachusetts Department of Education, with its regional offices throughout the Commonwealth, is the most appropriate and best suited agency to accomplish the technical tasks of instrument revision and participant training, and to provide general technical and logistical support to schools in the writing of objectives, analysis of relevant outcome measures, and the ongoing monitoring of development on the school factors. This is especially true given the current practice of Dr. Janice Weinman, Director of Assessment, and
others in the Bureau of Research and Assessment of using the regional offices as major vehicles in the assessment of education in Massachusetts. Financial support of the collaborative involvement of city schools in the process of self-improvement is provided for in the recent legislation amending Chapter 40 of the General Laws of the Commonwealth. The newly worded Section 4E of this law states that the department of education “through its regional offices shall assist and encourage cities . . . to enter into collaborative agreements . . . to conduct jointly educational programs and/or services which permit [them] to supplement or strengthen school programs and/or services.” (italics added) Clearly this legislation and the concern of the Commissioner of Education to provide meaningful support to schools make the regional offices ideal for the job of orchestrating the various aspects of applying the process it was the purpose of this study to refine and test.

Thus schools have available to them a process for self-examination, a regional structure for support in acquiring the necessary skills, and financial assistance through state legislation for carrying out the process. The remaining critical element can be supplied only by the schools themselves. It is commitment to improvement.

NOTES

Note 1 While some of the names of these factors match those of the Weber Study, in many cases the definitions are quite different. The intention was to focus even more than did Weber on the processes of delivering instructional services.

Note 2 Two schools were dropped from subsequent analysis because their factor ratings, taken as a set, were inconclusive.

Note 3 When we set aside coordination and individualization from consideration, this is not to suggest that these factors are necessarily irrelevant to a school’s success. Since all four types of schools in the study have in common low ratings on them, these factors do not contribute to our effort at contrast.
REFERENCES


Comments on Mace Report on Reading
(Use additional sheets if necessary.)

The Massachusetts Advisory Council on Education invites your comments on this report. Any statements you mail to the Council BEFORE APRIL 1, 1975, will be carefully considered by the Council in its review of the conclusions and recommendations of the study director.

The form to the right of this invitation may be clipped and used to record your comments. Attach additional sheets if necessary. Also, if you wish to appear before the Council to offer verbal testimony on this report, record your request on the form. Mail your comments and/or request before April 1, 1975, to:

Director of Research
Mass. Advisory Council on Education
182 Tremont Street
13th Floor
Boston, Mass. 02111

Thank you.

From (print name): ____________________________________________
Organization (if any):_________________________________________
Address: ____________________________________________________
Tel. __________ Date: ________________
Signature ____________________________________________________
## SOME OTHER RECENT REPORTS OF THE MASSACHUSETTS ADVISORY COUNCIL ON EDUCATION

<table>
<thead>
<tr>
<th>TITLE</th>
<th>AUTHOR</th>
<th>WHERE AVAILABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massachusetts Taxes: A Factual Guide to Future Action</td>
<td>Massachusetts Taxpayers Foundation, Inc.</td>
<td>MACE</td>
</tr>
<tr>
<td>Effectiveness, Efficiency and Equal Opportunity</td>
<td>Governor's Commission on School District Organization</td>
<td>MACE</td>
</tr>
<tr>
<td>Aid to Private Higher Education in Massachusetts</td>
<td>Frederick E. Terman</td>
<td>MACE</td>
</tr>
<tr>
<td>Higher Education in Massachusetts: A New Look at Some Major Policy Issues</td>
<td>Academy for Educational Development</td>
<td>Summary only</td>
</tr>
<tr>
<td>Strengthening the Alternative Post-Secondary Education System</td>
<td>George Nolfi and Valerie Nelson</td>
<td>MACE</td>
</tr>
<tr>
<td>Continuing and Part-Time Study in Massachusetts</td>
<td></td>
<td>4-page Summary Only</td>
</tr>
<tr>
<td>Something of Value (Summary) and Elementary Science Handbook</td>
<td>Office of Instructional Research and Evaluation Harvard University</td>
<td>MACE</td>
</tr>
<tr>
<td>The Here, Now and Tomorrow of Cable Television in Education</td>
<td>Creative Studies, Inc.</td>
<td>MACE</td>
</tr>
<tr>
<td>Modernizing School Governance for Educational Equality and Diversity</td>
<td>Paul W. Cook, Jr.</td>
<td>MACE</td>
</tr>
<tr>
<td>Massachusetts Schools Past, Present and Possible</td>
<td>Richard H. de Lone</td>
<td>MACE</td>
</tr>
<tr>
<td>Child Care in Massachusetts: The Public Responsibility</td>
<td>Richard R. Rowe</td>
<td>ERIC ED No 065-174 (full)</td>
</tr>
<tr>
<td>A Systems Approach for Massachusetts Schools: A Study of School Teaching Costs</td>
<td>Campbell, Aldrich &amp; Nulty</td>
<td>ERIC ED No.060 531 (full)</td>
</tr>
<tr>
<td>Continuing Education in Massachusetts State Programs for the 70's</td>
<td>Melvin Levin</td>
<td>D.C. Heath Publishing Co. Lexington, Mass.</td>
</tr>
<tr>
<td>Guidelines for Planning and Constructing Community Colleges</td>
<td>Bruce Dunsmore</td>
<td>ERIC ED No.034 390</td>
</tr>
<tr>
<td>Pupil Services for Massachusetts Schools</td>
<td>Gordon Liddle and Arthur Kroll</td>
<td>ERIC ED No.037 767</td>
</tr>
<tr>
<td>Take a Giant Step: Evaluation of Selected Aspects of Project 750</td>
<td>Herbert Hoffman</td>
<td>ERIC ED No.061 695</td>
</tr>
<tr>
<td>Teacher Certification and Evaluation in Massachusetts</td>
<td>Lindley J. Stiles</td>
<td>ERIC ED No.027 243</td>
</tr>
</tbody>
</table>