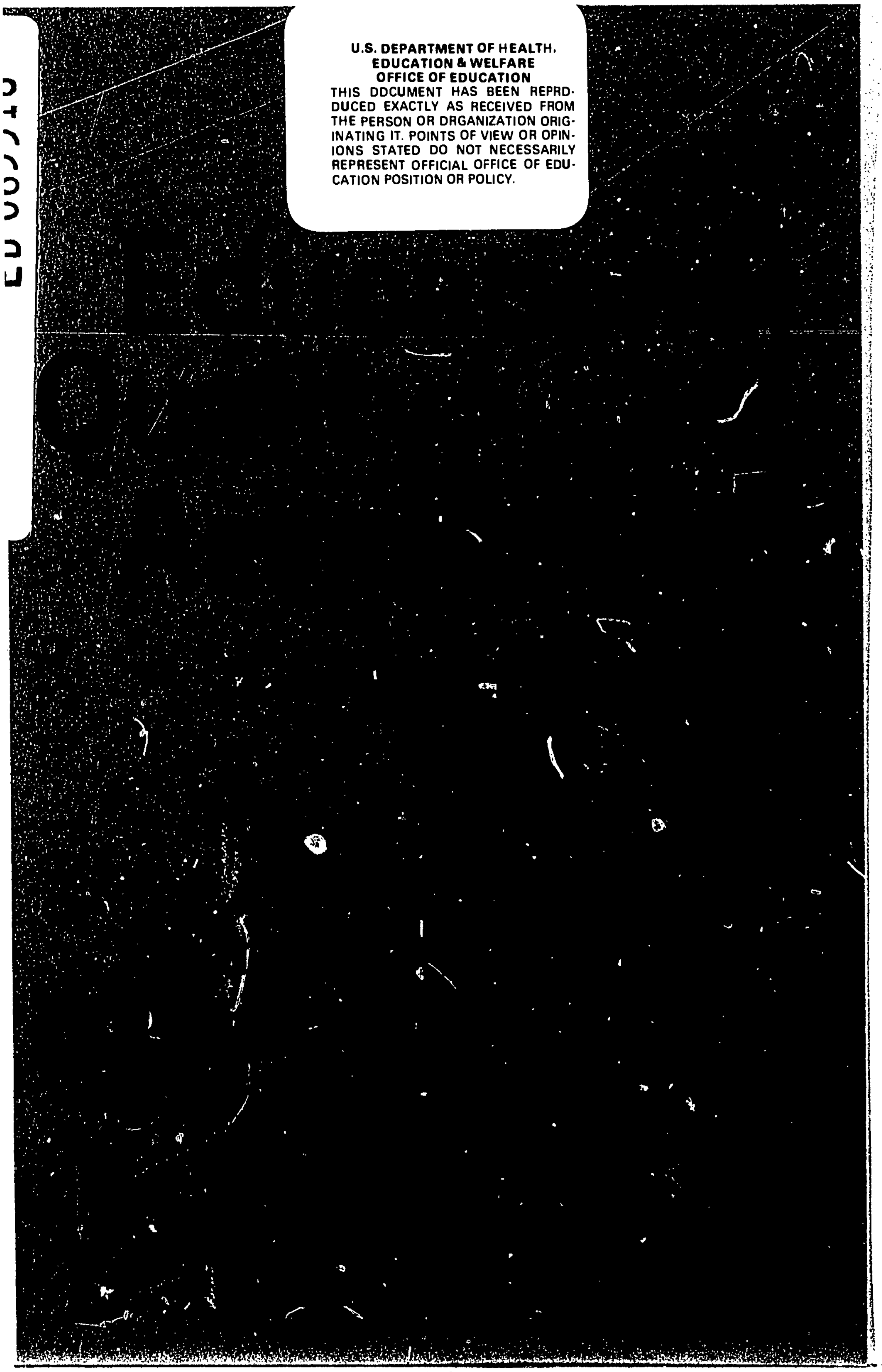
DOCUMENT RESUME

ED 065 516 TM 001 436

AUTHOR	Welsh, James
TITLE	Educational Quality Assessment. Pennsylvania Looks at Its Schools.
INSTITUTION	Pennsylvania State Dept. of Education, Harrisburg.
PUB DATE	71
NOTE	26p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Educational Objectives; *Educational Quality; *Evaluation Criteria; *Evaluation Methods; *Measurement Instruments; Predictor Variables; Questionnaires; State Programs; *Student Evaluation; Test Reliability; Test Validity
IDENTIFIERS	*Pennsylvania

ABSTRACT

The Pennsylvania Plan for Educational Quality Assessment is discussed as to its beginnings, its goals, its implementation, and its future. The plan aids schoolmen of the State in assessing the quality of education their schools are providing in relation to 10 goals, which propose to encourage the students to develop self-understanding, understanding of others, basic language and number skills, good citizenship, good health habits, their own creative potential, an appreciation of broad human achievement, a continuing interest in their own further learning, an ability to prepare themselves for productive vocations, and an ability in general to change within a changing world. In 1968, a pilot program (Phase I) was implemented to test a 10-goal measurement package to assure that each test was reliable and valid. The variable conditions were characteristics of students, of teachers, and of school and community. Phase II, carried out in 1969, used replies of students and teachers to questionnaires and data from school administrators and the Bureau of Statistics to develop standards for students' performance. Phase III, begun in 1970 and still underway, is the actual assessing of 111 school districts. A case study of one of the school districts is included with a chart of the percentile ranking of the schools on the 10 goals, as well as the actual score for each of the school's students. (For related documents, see TM 001 437-439.) (DB)



Educational
Quality
Assessment

Pennsylvania Looks At Its Schools

by James Welsh

Pennsylvania Department of Education
September 1971

Commonwealth of Pennsylvania
Milton J. Shapp, Governor

Department of Education
David H. Kurtzman, Secretary

Office of Basic Education
B. Anton Hess, Commissioner
Donald M. Carroll Jr.,
Assistant Commissioner for
Programs and Services

Bureau of Educational Quality Assessment
Thomas E. Kendig, Director

Pennsylvania Department of Education
Box 911
Harrisburg, Pa. 17126

Foreword

The Pennsylvania Department of Education has developed a plan for assessing the quality of schools in the Commonwealth. A number of department publications describe the procedures employed in the design of the Pennsylvania Plan (see Sections 1-6 of *Phase II Findings*).

James Welsh, a staff writer for the Washington *Evening Star*, provides the interested reader with a less technical view of the plan.

Contents

Pennsylvania Looks At Its Schools

Introduction

NOT long ago, in a mood of exasperation, the American Association of School Administrators stated that "no public institution in the world is assessed so frequently and critically as American education." At the same time, a faculty group at the Harvard Graduate School of Education was saying that "the improvement of the quality of education has always been hampered by our remarkable ignorance of what happens to young people as a result of the time and money expended on them in schools."

The two expressions of opinion are not nearly so contradictory as they might appear. In fact they are quite compatible, for they go to the heart of some of the knottiest questions to be found in American education today: How good are our schools? How do they compare with the promises that have so long been made? Are they doing the best they can? How is it possible to measure all this? And if we are to measure the performance of schools, how can we do so with sufficient meaning to generate whatever changes are necessary? Both the school administrators and the Harvard professors were saying in effect that frequent assessment and criticism of schools is one thing; meaningful assessment is another.

It is unlikely that either of those statements would have been made a short 10 to 15 years ago. Public schooling in America historically has been shrouded in faith and optimism. Until less than a decade ago, the promise and power of formal schooling were rarely questioned. Always the assumption was that with sufficient resources—enough classroom space, enough learning materials, enough trained teachers—the American dream could be realized: each school and school district performing smoothly to provide opportunity to all, to carry each child to his capacity in the learning skills and in civilized behavior. Sometime during the '60s the realization dawned that things were not all that simple. Catching up from the Depression and wartime restrictions, America had closed much of the gap in quantity of educational resources. But quality of educational performance lagged. It became apparent, for one thing, that millions of American children, especially the children of lower-income minority families, were not doing well in school, if

they stayed in school at all. The federal government, trying to correct the balance, moved in with programs and money, but with mixed results. The Coleman report, a massive social-science study that collected data from 600,000 pupils in schools throughout the country, concluded that in what they actually do, as contrasted to outside influences, the schools exert remarkably little impact on what becomes of these children. Meanwhile, as the decade progressed, the schools became a focal point for social protest and the aspirations of community groups.

People were saying to their schools: Do something! The schools, many of them unsure of their own performance, did not know exactly what to do. Their uncertainty, if anything, has compounded the frustration of the public. Call it part of the consumer movement, if you will. People—parents, the general public, even some of the student body, people in central cities, the suburbs, the countryside—have come to expect of their schools a greater measure of candor, and greater assurance of effort and results.

Out of all this, from the variety of forces at work in recent years, has come the theme of "accountability." It is a curious term. It means many things to many people. Basically, however, the idea it conveys is that the schools, or to be more precise, the professionals who run the schools, should be answerable for their product. Accountability demands that the educators, to a greater degree than heretofore, should be responsible for educational outcomes. After all, it could be argued, aren't the big Detroit auto firms held accountable for the performance of the cars they sell us? Why can't schools do the same? The parallel is not so simple, some might say. Indeed, it's not. A child—who he is and what he learns, values and aspires to—is not a precisely measurable piece of machinery. Two schools, moreover, can have markedly different raw material with which to work. And so there are sticky questions associated with trying to hold schools accountable. But for all that, the demands that schools become more product-responsible appear too strong to be thwarted. Accountability is here to stay, and the only question is how to make it operational.

Henry Dyer, one of the nation's foremost authorities on testing and evaluation, addressed this question in a recent article in the *Phi Delta Kappan*. At the level of the individual school, Dyer stated, accountability can be said to embrace three general principles:

1. The professional staff of a school is to be held collectively responsible for knowing as much as it can about the development of its pupils, and about the conditions that may be helping or hindering the pupil's development.
2. The professional staff is to be held collectively responsible for using this knowledge as best it can to maximize the development of its pupils toward certain clearly-defined goals.

3. The Board of Education has a corresponding responsibility to provide the means whereby the staff of each school can carry out its own responsibilities.

Dyer's breakdown of where responsibility lies, and how it can be exercised—know thyself and use the knowledge—could serve as a practical launching pad for the accountability principle. It can also serve as a point to introduce what this booklet is all about.

Enter the Pennsylvania Plan for Educational Quality Assessment. It is a program that will help every school in Pennsylvania to meet fully Dyer's first know-thyself principle. And it will serve as an unparalleled platform for schoolmen and school boards throughout the state to act upon the second two principles. It was designed at a time when the term "accountability" was unknown in the educational vocabulary. Now ready to go to work, the Pennsylvania Plan is tailormade for the age of accountability.

The Pennsylvania Plan

The Time Has Come

IN the spring of the year, teams of educators fan across Pennsylvania, armed with results of the Quality Education Assessment program, the first concrete results that can be of specific benefit to local schoolmen. They go into all the districts that requested to take part in the current phase of the program.

The state-directed teams hold intensive briefing sessions with local superintendents, principals and other administrators. On a school-by-school basis, the information unfolds. It tells each principal, along with his supervisors and supporting personnel, just where his school stands, what its strengths are, and what its weaknesses are.

It may tell him, for instance, that his 5th graders are doing better in basic word and number skills than most 5th graders across the state but that in comparison to similar schools, schools where similar sets of conditions prevail, those same 5th graders are not doing so well as should be expected.

Or it could tell the high school principal that in understanding of people from different cultural and ethnic groups, his junior-class students are about at the state norm and doing better than students in schools that can be called similar to his own.

The information to be given to local schoolmen goes far beyond that. It goes to other educational goals, such as self-understanding, creativity, and vocational readiness. It goes to those in-school and out-of-school conditions that play so big a part in a child's development and in the rationale of the Pennsylvania Plan. From family socio-economic levels to the educational background, experience and stability of a school's teaching staff, dozens of these conditions play a part in the educational climate of a school. Local educators no doubt are aware of them. But in the briefing sessions, the state-directed teams help the schoolmen go farther. They interpret the conditions at work in each school, showing where the school stands in relation to other schools and how important (or unimportant) each of these factors is to each of the performance goals the schools are trying to meet.

In brief, the Educational Quality Assessment program is first and foremost an instrument for self-understanding. It provides local school-

men what they need to know to assess the kind of job they have been doing so far. The information will be cheerful to some, alarming to others, a mix of good and bad news to still others. Either way, it should carry local school officials a long way towards determining what courses of action to take in the future.

To say all this, of course, is to raise a number of vital questions. What are the dimensions of this assessment program? Who set them, and why? On what basis can schools be judged as doing an excellent or a poor job? What are the limitations as well as its promises?

Here it is necessary to backtrack somewhat, to look at the Pennsylvania Plan as it first began and then took solid shape. It is necessary to explain the Pennsylvania Plan's philosophy and to define clearly both what it is intended to do and, equally important, what it is not intended to do.

Premises and Policy

The roots of the Pennsylvania Quality Assessment Plan lie in a provision the Legislature added to state law in August, 1963. It directed the State Board of Education to develop "an evaluation procedure designed to measure objectively the adequacy and efficiency of the educational programs offered by the public schools of the Commonwealth." This "evaluation procedure," said the law, should include "tests measuring the achievements and performance of students pursuing all of the various subjects and courses comprising the curricula." And the Legislature directed that the procedure should also enable each school district to appraise its own educational performance and, beyond that, to begin "without delay" the strengthening of its school programs.

It was a broad, challenging mandate. And infinitely easier said than done.

Not that the idea of assessing schools was completely new. It was an idea that had been considered often and at length by educational theorists, by other state education agencies, by private organizations (the Carnegie Foundation was one) that hoped to put the concept into motion. But always it was discussed timidly. For all its attractions, great obstacles remained.

There was, in the first place, the hard fact that no one so far had devised ways to measure exactly how well schools were doing what they were supposed to do. And even if it were assumed that the goal of the schools was little more than to teach mastery of skills and subject matter, was it possible to build into the assessment formula a recognition of the great diversity among schools, their students and the many variables of home and community life? It was on this point that suspicions grew among teachers and school administrators, those people who would possibly come under the gun in any assessment program. They

feared, understandably, an assessment formula that failed to take into account the complexities of education, a formula that would lead to snap, surface judgments of what they were doing. They argued, moreover, that some schools would hold a positive, inescapable advantage over other schools in any assessment program that might be devised. And the more they argued this point, the more they gave the lie to the old, fundamental assumption that schools are altogether the great leveler of opportunity. The wealthy suburban school, for instance, appears to have everything going for it—an educated, alert group of parents spurring their children on, the students themselves so often well motivated and reaching for razzle-dazzle academic performance, the finances of the school district ranging from adequate to handsome, the schools able easily to attract the best of the new young teachers. In contrast, the low-income urban schools labor under all the inner-city ills, represented most intensely by the great numbers of children who come to them ill-equipped to learn by ordinary methods. It is these schools, moreover, together with schools in the poor coal-valley towns and remote rural areas, that have the most difficult time attracting crack teachers and supporting personnel. How to reconcile all this? And finally, as one more obstacle to the development of an assessment plan, the argument was put forward that perhaps academic success is not the be-all and end-all of student performance. And once this was argued, the question arose as to exactly what it is that Americans expect their schools to accomplish.

This last question was one of the first settled in the shaping of the Pennsylvania Plan. Following the mandate of the Legislature, the State Board of Education named a committee on quality education headed by Mrs. Elizabeth Greenfield, chairman of the board's Committee for Basic Education. This group actively sought the advice of educational leaders across the country and, with aid of Educational Testing Service of Princeton, N. J., agreed on just what quality education is.

It wasn't long before the quality-education committee reached a basic conclusion: Academic success is important; but it is not the only goal of quality education. Nor is it the key goal, said the committee, for success in the traditional three R's and in other subject areas neither assures nor predicts self-direction, good citizenship, compassion for others, or much of anything else. The committee took the position that the schools should place high priority on their students' personal and social growth as well as their intellectual growth.

In March, 1965, the State Board adopted the Ten Goals of Quality Education that would serve as the basis for the assessment plan. In brief, the board ruled that Pennsylvania schools should lead children to develop self-understanding, understanding of others, basic language

and number skills, good citizenship, good health habits, their own creative potential, an appreciation of broad human achievement, a continuing interest in their own further learning, an ability to prepare themselves for productive vocations and an ability in general to change with a changing world. As the State Board adopted them and as they still stand, the 10 goals stipulate that quality education should:

- **Help every child acquire the greatest possible understanding of himself and an appreciation of his worthiness as a member of society.**
- **Help every child acquire understanding and appreciation of persons belonging to social, cultural and ethnic groups different from his own.**
- **Help every child acquire to the fullest extent possible for him, mastery of the basic skills in the use of words and numbers.**
- **Help every child acquire a positive attitude toward the learning process.**
- **Help every child acquire the habits and attitudes associated with responsible citizenship.**
- **Help every child acquire good health habits and an understanding of the conditions necessary for the maintaining of physical and emotional well-being.**
- **Give every child opportunity and encouragement to be creative in one or more fields of endeavor.**
- **Help every child understand the opportunities open to him for preparing himself for a productive life and should enable him to take full advantage of these opportunities.**
- **Help every child to understand and appreciate as much as he can of human achievement in the natural sciences, the social sciences, the humanities and the arts.**
- **Help every child to prepare for a world of rapid change and unforeseeable demands in which continuing education throughout his adult life should be a normal expectation.**

And so, in one all important area, the State Board settled decisively what the plan is and what it isn't. To reiterate: It is a plan that attempts to measure the success of the schools in teaching the three R's, but it is not limited to that. If the Pennsylvania Plan means anything, it is that schoolmen at both the state and local level are called

on to look at the *whole* child, and at how the schools are doing in developing the *whole* child.

This was both start and foundation. The bulk of the Pennsylvania Plan had yet to be fleshed out, and important questions remained to be answered. But over a period of time, state officials, continuing to work with education and civic leaders from throughout the Commonwealth, worked out basic decisions and assumptions that comprise the working philosophy of the program. It is a philosophy, again, that can be expressed by saying what is, and is not, true of the Pennsylvania Plan.

It *is not* a plan leading merely to a blanket statement of how the overall Pennsylvania establishment is doing. Here it differs from the federally-financed National Assessment of Educational Progress, which so far has covered a small sample of pupils representing broad regions of the nation.

Neither, on the other hand, is it a plan designed to show how Susie or Johnny or any individual student is doing. It *is* a plan that focuses on the individual school as a unit, and it is designed to lead the people who run each of these schools to understand how well or how badly their programs are serving all the Susies and Johnnies who attend them.

It *is not* a plan that assumes education proceeds in a vacuum, or that individual schools, any more than individual pupils, begin with a clean slate.

It *is* a plan that takes into direct account the many conditions, some within the school, some outside the school, that influence for better or worse how teachers are able to teach and how pupils are able to develop. To put it another way, it is a plan that lets schools know how well they're doing in light of the conditions under which they operate.

It *is not* a program through which officials from Harrisburg tell local schoolmen what to teach, what not to teach, what to emphasize or play down in their school programs.

It *is* a plan mandated by the state, a plan hinged to goals of quality education established by a state agency and a plan in which the state plays a strong, specific role. But in all state-local dealings, this role is one of consultation only.

Educational Research at Work

With the establishment of the Bureau of Educational Quality Assessment in mid-1967, a most important phase of the Pennsylvania Plan was ready to begin. This was the hard-slogging work of translating the mandate of the Legislature and the wishes of the State Board of Education into a workable plan that could serve local schoolmen.

The bureau's staff of education researchers knew pretty much what they had to do. Given the Ten Goals of Quality Education, they had to construct for each of them a set of measuring tools that would determine the standing of Pennsylvania schools in meeting those goals. Beyond that, in putting together what became known as the "assessment model," they were faced with the task of identifying those environmental conditions, or variables, at work within schools, related to student performance on each of the 10 goals.

First priority went to measurement. Here the bureau staff found its job very easy in some respects, very difficult in others. For example, measuring achievement on Goal III—basic skills in the use of words and numbers—presented no problems. American education long has been concerned with gauging academic achievement. Suitable measuring devices in the form of standardized tests were readily available. For most of the other goals it was a different story. Few satisfactory measures were available. The bureau staff investigated the literature and solicited help from test publishers and researchers elsewhere who were interested in the science of measurment. On a few goals—for example, Goal X, the ability to change with a changing world—the staff was in virtually uncharted territory. It had to brainstorm its own items for a questionnaire.

By the spring of 1968, a 10-goal measurement package was ready for pilot testing. This was done at 100 schools throughout the state, with some 2,700 pupils taking part. These pupils were at two grade levels, the 5th and 11th. Fifth grade, it was decided, was a good place to measure the elementary-school experience. For the high-school level, tests at the 12th grade level were ruled out because seniors already have so many demands on their time. Testing at 10th grade would reflect but a small part of the senior high school experience. The 11th grade was selected as the most suitable for testing. Junior high schools also were important to assess, the bureau recognized, but in part because of financial limitations at the state level, it was decided to bring in these schools later in the plan's development.

The 1968 pilot program was a test not of the schools but of the tests themselves and each of the items that made up these questionnaires. One question, answered 95 per cent in an unsatisfactory way, might be thrown out as too difficult. Another item, left unanswered by half the pupils, might be considered too vague and thus thrown out or reworked. So it went.

Throughout, the aim of the state-level researchers was to see that each of the tests was reliable—that is, each test item consistent with the others—and that each of the tests was valid—that is, each of the tests and its separate items measuring truly what it was supposed to

measure. All this was accomplished, in part through painstaking statistical analysis, in part through trial-and-error studies of various hypotheses that were formulated by the bureau staff. As an example of the latter, it is interesting to look at what was done about testing pupils for creative tendencies, or Goal VII. In one study, 5th graders were administered the test called Assessment of Creative Tendencies. At the same time, on the assumption that curiosity has much to do with creative tendencies, the teachers of these pupils filled out a curiosity behavior profile developed by a Johns Hopkins University researcher. The results were that the students' self-reports and the teachers' judgment of them were positively related. In yet another study, two groups of high school students were given two separate tests, one a "creative potential" test, the other a "creative output" test. Group 1 was limited to those 11th graders who already had been judged to be creative or who had won prizes in various art or science contests. It was hypothesized that Group 1 would score significantly higher on the two tests, and this proved to be the case.

Meanwhile, as the tests were refined and refined again, the Bureau of Educational Quality Assessment was also busy considering the dozens of variable conditions that might influence student performance. These, it was soon recognized, fell into three broad categories:

1. Characteristics of students. They include, it was hypothesized, the occupational and educational levels of parents. For secondary students, it was decided, they might also include the personal values and mores of the student.
2. Characteristics of teachers. These might include the background of the teachers' parents, the teachers' level of training, their age, experience and salary, and their attitudes toward their jobs. For example, it was hypothesized that it may well influence student performance whether teachers believe that their professional recognition and advancement comes from their own merit or their ability to gain an "in" with the principal or central office.
3. School and community characteristics. These, the bureau staff decided, might include the staff-pupil ratio, the counselor-pupil ratio, the number of library books available to each pupil, the financial ability and spending of the school district, the type of community in which the school was located, the degree to which students were exposed to youngsters of different racial origin and the school's dropout rate.

There might have been other school variables considered—the curriculum of each school and the way the curriculum was handled and taught. But these are factors extremely difficult to pin down to the

kind of specific items that can be answered and collected on a mass scale. Difficult to bite on, difficult to chew. And so the bureau decided, at least for a time, to settle for those school conditions that could be pinpointed with reasonable clarity.

More than 50 potential student, teacher and school variables were selected for consideration in the Phase I pilot program of the Pennsylvania Plan. From student questionnaires, from teacher questionnaires, from state school records, the bureau gathered information on each of these conditions. Each of them was tested with the most modern statistical techniques. The obvious interrelationships between many of them were analyzed. The ultimate goal was to find the exact relevance of these conditions to the performance of students in each of the 10 quality education goals, and from there to determine what the likely performance scores might be for students in each school in Pennsylvania.

Putting It All Together

The 1968 pilot testing could carry the Pennsylvania Quality Assessment Plan only so far. It remained for the state Bureau of Educational Quality Assessment, using what it learned in the pilot project, to conduct a much bigger data collecting program in order to establish normative standards for the measuring devices for each of the ten goals of quality education. It is one thing to write items for a questionnaire but quite another to know what to expect as an "average," "high" or "low" score. The purpose of Phase II was the development of such standards.

Phase II of the plan got underway in the fall of 1969. It was, indeed, bigger in every way than Phase I. Taking part were some 20,000 5th graders and 17,000 11th graders (approximately 10 per cent of the state's school population in those grades), together with nearly 2,000 of their teachers.

For each of the pupils, it was an unusual assignment. The pupil was asked to be frank and honest in filling out a questionnaire containing several hundred items. In return, the pupil was promised the questionnaire would be confidential, not to be shown to teachers or administrators. Some of the questions bore directly on those variable conditions with which the pupil was familiar such as father's occupation. The student was asked to react to statements such as: there is little chance for a person like me to succeed in life (Goal I); I would like sitting next to a pupil of a different color, or different religion, or poorer family (Goal II); I would like to quit school soon (Goal IV); it's okay to break a school rule if everyone else is breaking it (Goal V); I have little or no idea what working for a living will be like (Goal VIII); I think most works of art are too difficult to understand (Goal IX).

The teachers, also assured of confidentiality, filled out a 76-item questionnaire, the items ranging from their own backgrounds to what satisfaction they get from their jobs. In addition, the state bureau compiled other data from school administrators and the Department of Education's Bureau of Statistics.

Into the computers the raw data went. Out came the basis for telling the administrators of each school what they could expect in their students' performance on each of the 10 goals. It wasn't all that simple, of course. Computers can be of benefit only in light of the analytical techniques fed to them. In the Pennsylvania Plan, the major statistical tool used was regression analysis, a technique that permits the researcher to predict the most likely average achievement score from the known information on such other variables as pupil, home, school and community conditions. For each of the 10 quality education goals, it became possible to construct a "regression equation" showing which condition variables contribute to prediction of the performance scores. It also became possible to show how much or how little these conditions relate to one another and, finally, to show how much the combination of conditions accounts for the achievement scores on each of the 10 goals.

At the outset it was stipulated that to be considered for any of the goals at either grade level, an environmental condition must contribute at least one per cent to the ups and downs of performance scores. Not surprisingly, some of the conditions—such as the occupation of the student's father and the education of the student's mother—proved highly related, at least on some of the goals. Other conditions, selected early, proved to be inconsequential. An example is the educational level of the teacher's mother. Still other condition variables were found to interact considerably with one another. An example here is teacher experience and teacher salary. All together, of the more than 50 conditions selected for consideration in Phase II, 45 were retained in the 11th grade and 39 in the 5th grade.

To show how the bureau's analysis worked, and what it means, it might be helpful to look at one illustration.

Consider Goal I, self-understanding. When all the data for 5th graders was processed, it was discovered that five conditions predicted performance. They were: father's occupation, housing conditions in the community, teacher stability, teacher experience and school subsidy per student. All together, these five conditions were found to account for 23 per cent of the variability of the pupil scores on self-understanding. These results could be called both disappointing and hopeful. They are disappointing in that the five influential conditions are difficult or impossible to change. It is not within the power of the schools to improve

a father's occupation or a community's housing stock. But the results also can be interpreted this way: If 23 per cent of a 5th grade pupil's performance in understanding himself and assessing his or her own worth is due to change-resistant conditions, the remaining 77 per cent is due to other, as yet unidentified factors. These factors, possibly including classroom practice and parent attitude, may be open to improvement by teachers and administrators.

In grade 11, the analysis of Goal I produeed strikingly different results. Here 14 separate conditions were found to be at work. All together they aecounted for 61 per cent of the student scoring on this goal. Some of the conditions, as with the 5th grade, resist change. But others are more open to manipulation. The single most related variable was what the teachers believed to be the best way to earn professional recognition. It indicates that when teachers believe reeognition comes from profcssional merit, not politics and conniving, eonditions exist that are associated with positive student self-concept. Other conditions that applied to performance by 11th graders on Goal I and that are open to change are teacher salary, the school district's tax effort, the accessibility of counselors and the counselor-pupil ratio.

Some Larger Findings

Although the major focus of the Pennsylvania Plan was to provide information of use to those running the individual school, analysis of the data from the Phase II survey revealed a number of overall findings that should interest anyone concerned with public education.

A major revelation was that, while not contradicting the earlier Coleman report, the Pennsylvania Plan goes well beyond it to show that in many areas of pupil achievement the schools are in a position to make many changes that will improve performance. The Coleman report was limited to academic achievement, which in the Pennsylvania Plan is represented by Goal III. It concluded that a very high percentage of student output, or performance, is associated with environmental conditions. Similarly, in analyzing the Goal III data, the Pennsylvania researchers found that for 5th graders, about 77 per cent of performance variation and for 11th graders about 87 per eent of performance variation can be attributed to environmental eonditions.

All together, it was discovered that 49 per cent of performance on all goals in both grades could be attributed to various environmental conditions, some that can be manipulated, some not. That leaves slightly more than half the performance level of students due to variables that may be influenced positively by improved curricula, teaching methods and classroom practices.

In considering specific goals, the researchers ran across other findings. On Goal II, it was discovered that a high degree of interracial exposure did not positively influence understanding of others. In fact it was a negative factor, although to a very slight degree.

As some of the preceding figures suggest, environmental conditions play a much bigger role in the performance of students at the 11th grade level than they do at the 5th grade. More variable conditions are at work, and on every goal they account for a considerably higher percentage of the performance scores. The researchers also discovered student development does proceed as children move through school. This was shown most graphically on three goals in which the tests were identical for 5th and 11th grades. On Goal I, self-understanding, Goal V, citizenship, and Goal X, preparing for a changing world, the performance of 11th graders was significantly higher than that of 5th graders.

The Current Phase

With the completion of Phase II, Pennsylvania had its quality assessment plan and normative standards for its measuring instruments. The only missing ingredient was schools to assess. In the fall of 1970, the Bureau of Educational Quality Assessment launched the Plan's Phase III, which is still underway. The state in effect asked for volunteers, and in a heartening response, more than half of Pennsylvania's 579 school districts asked for assessment. The 111 districts selected to take part in the first Phase III included nearly 600 schools. Once again 5th and 11th graders were the targets, more than 50,000 of them.

These students were tested and questioned as other youngsters had been the year before. So were their teachers. This time around, however, with the analysis of previous data in hand, the Pennsylvania researchers were able to pin down the expected achievement scores for each school according to the environmental conditions prevailing there. Then they could compare the *actual* scores, as revealed by the Phase III tests, to the *expected* scores predicted by the bureau with the Phase II data in hand. With that accomplished, they were in a position to tell the administrator of each school whether, on any of the goals, and in comparison to schools where a set of similar conditions exists, his or her 5th and 11th graders were performing as expected, or better, or worse.

At this point, it would be well to look at one specific school district to see how the plan operates.

Sunrise Area School District: A Case Study

CALL it Sunrise Area School District. Think of its superintendent as Henry Brown. Both the school district and the man in charge are real. Only the names have been changed.

The school district is fairly small, less than 20,000 residents. It is composed of one small city of less than 10,000 people and several neighboring boroughs and townships. It contains six elementary schools, a junior high school, and a senior high school. Sunrise Area School District is in a mountainous area of Pennsylvania. Yet it is not very far from industrialized cities. It used to be a place with an independent economy. Now, like so many communities in Pennsylvania, it is a place where rural life proceeds as usual but where many people in both town and country get up early in the morning, commute to jobs in cities as far as 50 miles away, and drive home at night. Some of these people make good money, some not so much. The children who attend the area schools therefore represent a mix of the fairly affluent, the working class, and stable farm families.

"I am very strong for evaluation," Superintendent Brown was saying recently. "I have looked for the best possible evaluation procedures in every school district where I've worked."

Mr. Brown was explaining why he voluntecred his school district early to take part in Phase III of the Pennsylvania Plan. As the data from the Phase III test-survey flowed into and out of the computers, two of the elementary schools in Sunrise Area School District were among the first for which full information was available. Mr. Brown, in the sessions to follow with personnel from Harrisburg, got all the evaluation he ever bargained for.

One elementary school in Sunrise Area School District is located in the district's one small city. It can be called the town school. In the survey and testing of the school's 5th graders, it was shown that a fairly high family socioeconomic level prevailed, that the level of the father's occupation was higher than the state average. Their teachers tended to come from the local area, to be fairly young, educated at state colleges, making average salary, believing for the most part their professional recognition comes more from personal relationships than work and merit.

A second elementary school is located in a farming area. Call it the rural school. The family socioeconomic status is low, mother's education low, father's occupation lower than average. The teachers are local products, nearly all older, with many years of service and high salaries. They believe their professional recognition came largely through merit. As with the town school, it is part of a school district that makes a lower-than-average tax effort to support education, that provides little in the way of extras such as library books.

All this was explained to Mr. Brown and his fellow administrators. The state interpreting team showed them by a series of charts just where the two schools stood in relation to other state schools on the many student-home-teacher-school conditions that influence student performance. Then the interpreting team unveiled, for each school, another chart that was to have great impact. The chart was arranged in 10 vertical columns, each representing one of the 10 goals of quality education. The Sunrise Area school officials could look down each column to see how their students had scored on that particular goal. The actual score for the school's students was marked by a red symbol shaped like a schoolhouse. Where this symbol was placed in each column, the officials could see the 5th graders' actual score on that particular goal. The higher the symbol was in the column, the higher the score. The officials could look to the left on the chart to see how any particular score fell into a percentile rank for 5th graders across the state. A percentile rank of 70 for one school would mean that the students there scored higher than 70 per cent of all schools in the sample, lower than the other 30 per cent. The chart revealed an even more essential comparison. Within the column for each goal were to be found two blue horizontal lines, or "prediction boundaries." One might be drawn through a score at the 60th percentile rank, the other through a score at the 30th percentile rank. Within these two boundaries was where the school's 5th grade test scores could be expected to fall, given the conditions under which the school operated. Regression analysis in Phase II of the Plan had revealed that, on the average, students in schools with similar environmental conditions were scoring in that range.

The charts that were shown to Sunrise Area schoolmen disclosed a fascinating mix of scores, both for the district's town school and its rural school.

Looking first at the rural school chart on the following page, the Sunrise Area officials could see that on Goal I, self-understanding, the red schoolhouse symbol fell at about the 50th percentile mark. It meant the rural school's 5th graders were near the center of the entire Phase II state sample, half the sample scoring above them, half below. The "prediction boundaries" for this goal were marked at the 88th percentile and the 28th percentile, and this meant the Sunrise Area rural 5th

PENNSYLVANIA SCHOOL NORMS-GRADE 5

DEBE-222 (10-70)

PERCENTILE DISTRIBUTION	Percentile Rank	INSTRUMENT										Percentile Rank
		SELF UNDERSTANDING	UNDERSTANDING OTHERS	BASIC SKILLS	INTEREST IN SCHOOL	CITIZENSHIP	HEALTH HABITS	CREATIVITY	VOCATIONAL DEVELOPMENT	APPRECIATING HUMAN ACCOMPLISHMENT	PREPARING FOR A CHANGING WORLD	
	Above											Above
	95	90.95	34.75	556	62.75	[illegible]	[illegible]	145.67	62.57	54.21	108.80	95
	90	89.94	34.23	546	61.88	168.65	31.16	144.46	61.87	53.76	106.52	90
	85	89.19	33.82	537	61.15	167.22	30.54	143.28	61.42	[illegible]	105.07	85
	80	88.74	33.59	531	60.70	166.37	30.23	142.34	61.14	53.00	104.19	80
	75	88.32	33.36	526	60.25	165.53	29.91	[illegible]	60.87	52.66	103.31	75
	70	87.94	33.12	520	59.82	164.69	29.59	140.62	60.63	52.47	102.63	70
	65	87.64	32.92	516	[illegible]	163.85	29.29	139.88	60.46	52.31	102.03	65
	60	87.35	32.73	512	59.21	163.03	29.02	139.20	60.29	52.16	101.43	60
	55	[illegible]	32.54	508	58.90	162.21	28.75	138.53	60.12	52.00	100.84	55
	50	86.78	32.34	504	58.60	161.39	28.48	137.85	59.94	51.85	100.34	50
	45	86.50	32.17	[illegible]	58.28	160.57	28.21	137.17	59.76	51.69	99.85	45
	40	86.22	32.00	495	57.96	159.40	27.85	136.49	59.57	51.52	99.36	40
	35	85.94	31.83	489	57.64	158.01	27.48	135.80	59.38	51.36	98.87	35
	30	85.65	31.66	483	[illegible]	156.62	27.11	135.12	[illegible]	51.19	98.37	30
	25	85.37	[illegible]	478	56.96	155.12	26.67	134.18	58.95	51.03	97.81	25
	20	85.09	31.12	472	56.43	153.61	26.19	133.17	58.63	50.73	97.26	20
	15	84.80	30.80	[illegible]	55.89	151.79	25.68	132.09	58.32	50.42	96.70	15
	10	84.20	30.28	455	54.98	149.57	24.91	130.56	57.81	50.03	96.03	10
	5	83.31	29.64	440	53.36	146.40	23.73	128.06	57.18	49.46	[illegible]	5
	Below											Below
MEAN		86.90	32.33	501	58.45	160.25	28.20	137.64	59.90	51.80	100.76	
STANDARD DEVIATION		2.21	1.56	035	2.78	7.49	2.44	5.32	1.60	1.60	4.57	

graders scored within the expected range, given the set of conditions prevailing at their school. On Goal II, understanding others, the rural 5th graders scored lower, at the 25th percentile mark, meaning only 25 per cent of schools across the state scored lower. Even so, this score also fell within the "prediction boundaries" for that school.

Then came an eye-opener for the local schoolmen. On Goal III, basic skills, the "prediction boundaries" revealed the rural 5th graders, compared to pupils in schools with similar conditions, should have scored somewhere in the 45th to 83rd percentile range. But the red schoolhouse symbol showed they actually scored at the 15th percentile mark, much lower than expected.

On Goal IV, interest in school, the 5th graders did about as well as expected. On both Goals V and VI, citizenship and health habits, they did much better, scoring up at the 95th percentile mark, above the prediction boundaries. Their scores on Goals VII and VIII, creativity and vocational development, were within the expected ranges. The score on Goal IX, appreciation of human accomplishment, was at the 85th percentile, at the top of the range of score expectation for that school. But in another turnaround, the score on Goal X, preparing for a changing world, was below the 5th percentile and lower than expected.

Much of the same high-low scoring phenomenon was true of the district's town school 5th graders. They were down at the bottom, the 5th percentile, in understanding of others, a good bit lower than the scores of similar schools. They were at the 50th percentile on the Goal III basic skills, but once again, this was much lower than they should have scored, since the prediction boundaries showed students in similar schools to be between the 70th and 95th percentile marks. In health habits, but not citizenship, the town 5th graders were at the top of the chart, above the range of expectation. In creativity and vocational development, they were also high, near the top of the expected scoring range. In ability to prepare for a changing world, they scored higher than the rural children, at the 50th percentile mark and well within the expected range of scores.

The Sunrise Area school officials were given additional material to analyze. For instance, they obtained key items on the tests, showing how each school's 5th graders scored on those questions. For further analysis, they obtained, for each goal, a break-down of the school test scores showing the distribution of students in various percentile ranges compared to students statewide and in similar schools.

Not long after the presentation, Mr. Brown sat back to reflect on what it all meant. "It has certainly given us a great deal to think about," he said.

To Mr. Brown and his colleagues in the school district, the real shocker was that in both the town and the rural elementary schools,

their 5th graders scored so much lower on Goal III than expected, lower than students in schools with similar conditions.

"We were amazed at the basic skills scores," he said. "We knew we had some problems. Previous testing had shown some of that. But we didn't know how deep these problems are. We didn't know we were worse off than similar schools."

Some of the other results were less surprising to Mr. Brown. The rural school's high scores on citizenship, for example, he attributed to the township's stable population, its traditionalist teachers and a general conformity of view. For much the same reason, he could see why on Goal X, ability to prepare for a changing world, his rural 5th graders scored so much lower than the city youngsters.

Even though the results were distressing at some points, Mr. Brown is glad he put his school district through the assessment ropes. "It's a good program," he said. "It's a vehicle for us. It gives us the motivation to really study our program."

Mr. Brown intends to take up the state's offer of sending in personnel to consult on program changes that might be called for. He recognizes there is no magic solution to his school district's problem of raising student performance. But he believes a combination of state and local expertise might help us considerably.

"Our first priority will be in the basic skills area," he said. "After that, we'll select from the other goal areas."

The Sunrise Area superintendent concluded: "There's a great deal here to analyze. We've got a lot to do."

The Pennsylvania Plan: Its Future

THE payoff period for the Pennsylvania Educational Quality Assessment plan has begun. But much is yet to come, both in the near future and in the next several years. Together with cooperating state agencies, the Bureau of Educational Quality Assessment will be moving on many fronts at once.

As noted before, teams of state personnel will be busy in the coming months interpreting the Phase III data to administrators of the participating Phase III school districts. Also as noted before, these interpreters will not make value judgments for the local schoolmen. Their job is to make sure the local administrators understand what the Phase III information means. It is also to get those administrators to a point where they can make their own analysis of where they are and what possibly can be done to improve their students' performance. The Phase III data for individual schools, incidentally, will not be made public by the state. But if the local school board and administrators choose to do so, it is perfectly acceptable to the state. In fact, it can be reasoned that making the information accessible to press and public will help the cause of accountability.

Once they sift through the information given them local schoolmen will have the option of asking for more outside help. It is probable, as in Mr. Brown's case, that many of them will plan to do so. What can be done to assist Mr. Brown in his search to improve his schools? For one thing, the state Bureau of Curriculum Development and Evaluation is prepared to consult with local districts. Its staffers will help these schoolmen identify those goals on which paths to probable progress is clear. And they can help chart these paths—possibly a different allocation of local personnel, or a new classroom practice, or selected spending increases. On some goals, the way to improve performance won't be easy. Others may hold more promise.

Meanwhile, the researchers at the state level intend a great deal of follow-up analysis of the Phase III data. They will watch for schools where children performed significantly above, or significantly below, the range of scores predicted for that school. They intend to go into some of these schools, to observe what goes on there, to try isolating those factors and practices that might account for the significantly

different scores. The result could well be the compilation of a set of suggested methods and practices that other schools might follow in trying for higher performance on any of the quality education goals.

Meanwhile, too, the bureau is moving to expand the Pennsylvania Plan to other grades. Fall 1971 begins the Phase I testing of a group of 7th graders. In the fall of 1972 will come Phase II for the state's 7th graders. In the fall of 1973 will come Phase III for the 7th grade and also Phase I for the 3rd grade. Eventually, the Pennsylvania Plan will provide assessment for five grade levels—3, 5, 7, 9 and 11. Eventually, too, as required by law, all school districts in the state will undergo quality assessment.

At this point, the Pennsylvania Plan can be said to offer a mix of hope and caution. It would be misleading to say it provides all the answers a school district needs, or any foolproof formula to get those answers. At the same time it has proved so far that all is not lost in a sea of socioeconomic or other difficult-to-change outside conditions. It has proved that what goes on in the schools can make a very real difference. It offers schools an unparalleled vantage point for looking at themselves and assessing both what has been done so far and what needs to be done. It is giving the public a new basis, within reasonable bounds, to hold its schools more accountable for what they do.