

*MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM:
HISTORY AND DEVELOPMENT*

*Edward D. Roeber
Michigan Department of Education
May 1987*

*Prepared Under Contract For
The Office of Technology Assessment
Congress of the United States*

*Michigan Educational Assessment Program:
History and Development*

Introduction

During the early- and mid- 1960s, growing concern about the educational attainments of the nation's children and youth and rising costs of education combined to create a new concept in education — accountability. Rather than being solely concerned whether our children could read or whether the best college or university would admit our sons and daughters, we began to ask ourselves more fundamental questions about our public schools. While people looked to public schools to further social advancement and stressed the importance of a good education in finding a rewarding job and attaining the “good life,” serious questions about the quality of our schools were being raised.

Increasing concern over the products of schooling was natural. We asked ourselves: what can students do? Surprisingly, little information was available. Although local testing programs had been around for years, little data was available about students across Michigan. This lack of information led to the development of a state assessment program in Michigan.

The Creation of the Michigan Model

By State Board action and request, funds were provided in fiscal year 1969 to begin a statewide program (for implementation by the end of January 1970) to conduct an annual testing of all fourth and seventh graders. Without adequate time to create the measures to be used and hardly time to decide what measures could be used, the Michigan Department of Education (MDE) contracted with Educational Testing Service to develop the first tests. Measures in mathematics, reading, mechanics of written expression, word relationships (a hybrid “aptitude” measure), a socioeconomic status

(SES) scale and an attitude scale were prepared. All of these measures were norm-referenced. Data on school buildings, districts and the state as a whole would be released to school district personnel only; public release of data would not occur, by promise of MDE. While district and school norms were prepared and percentile ranks released, none of the data was made "public."

Obviously, such a large-scale program could not be implemented without controversy and if the state assessment program was strong on anything, it was strong on generating controversy! Teachers disliked the achievement measures. Low scoring districts disliked the percentile ranks. Parents and students were offended by the questions in the SES measure and turned off by the attitude scales. Administrators were defensive about potentially unfair comparisons, while teachers were worried about evaluation based on these test results.

Despite (or perhaps because of) this controversy, the program was continued through legislative mandate and funding (Public Act 38 of 1970). The second year of the program was even more controversial. Several large cities threatened to withhold their answer sheets from scoring if they were required to administer the SES and attitude scales.

The clincher came on Valentine's Day, 1971, when the State Superintendent, at a news conference well attended by the press, released a report of achievement results for every school district in Michigan. Although this seemed contradictory to the earlier promise of not releasing the results, the Department had been required by a state Attorney General's opinion not only to make the data public, but also to publish the data and disseminate it. Several newspapers in the state published the assessment scores; one paper (with statewide circulation) did so for all Michigan districts. That infamous day became known within MDE as the St. Valentine's Day Massacre: educator outrage and concern about the program reached its peak.

Efforts were begun in 1971 to work with mathematics and communication skills educators to refine the tests. For the first time, Michigan educators were writing test items. Items written by teachers appeared to be better measures of achievement of Michigan students and were better accepted. At the same time, two other fundamental changes occurred: 1) a model was developed that tied the state assessment program to statewide curriculum improvement and 2) the seeds of a new program were sowed.

In 1971, the six-step accountability was proposed and adopted by the State Board of Education in 1972. The model called for 1) the development of Common Goals, 2) the statement of explicit student expectations in the form of student performance objectives, 3) a needs assessment to determine specific student needs, 4) an analysis and modification of the instructional system where student needs are shown to exist, 5) an evaluation of the effectiveness of these changes in meeting students' needs, and 6) recommendations for future action.

As the efforts to develop the Accountability Model and the components of it were under way, the Assessment Program continued the annual administration of the norm-referenced tests in 1972 and 1973. Due to the continued controversy surrounding their use, the attitude scale and SES inventory were withdrawn.

Substantial item tryouts were held in 1971-72 to validate the teacher-written items for the achievement tests. New items were substituted into the achievement tests in 1972-73, marking the introduction of the first "nonprofessional-internal-writer" items in Michigan.

The final year of normative testing drew to a close in January 1973, with barely a whimper, for a far more exciting and innovative program lay ahead — the first use of objective-referenced tests on a statewide basis. 1972-73 was overshadowed by the new program.

Michigan's New Assessment Program

During 1971 and 1972, as the controversy surrounding the Assessment Program continued and as the misuses of the norm-referenced data mounted, a basic shift in the Assessment Program occurred. A decision was made by the State Superintendent and the State Board of Education to shift the Assessment Program to the measurement of objectives developed in Michigan. Tests would be developed for the minimum performance objectives in mathematics and reading.

Based on the previous successful experience of using classroom teachers to write and try out test items, a test development program was begun in 1972 with five school districts representative of the state, as well as a testing company to edit the items. Teachers, after receiving training in item writing, worked for several months to produce the needed items. The testing company then was responsible for editing a selection of the items and putting them together in tryout packages. The items were tried out. After tryouts, extensive reviews of the objectives and test items were conducted and the final fourth and seventh grade tests were assembled.

In the fall of 1973, the first objective-referenced assessment of students was conducted in Michigan. This was the first use of an objective-referenced test on such a wide-scale basis. Results were reported back for each student (and the student's parents), classroom teachers, building principals and central office staff. Considerable emphasis was placed on using the results to provide remedial instruction to the students tested, using the results to review and improve the school curricula, and reporting results to the parents, school board and the public, via the news media. The results were not used in promotion/retention decisions about students, nor were they tied in any way to high school graduation. The data have been used, though, as the basis for allocating state-level compensatory education funds (around \$30 million per year) to local districts. The switch from norm-referenced to objective-referenced tests was not without problems, however. First, the objective-referenced tests were longer, with

students needing up to four or five hours to finish the test. Second, because the tests were untimed, some educators did not know what to do with students who finished early. Third, the concept of a “minimal” objective was new — could all students attain all of the “minimal” performance objectives? Finally, there was concern over proper use of the results. Because of the number of performance objectives tested, and because of the decision to return results in a form useful to classroom teachers, assistance had to be provided in person and in writing to help teachers and administrators throughout the state to understand what the test data could (and could not) be used for.

Expansion of the Michigan Educational Assessment Program (MEAP)

When the mathematics and reading performance objectives were first written, they were divided into three sets: grades 1-3 (tested at grade 4), grades 4-6 (tested at grade 7) and grades 7-9. Tenth grade assessment was seen as a logical extension of the fourth and seventh grade program. Test development began in 1974 and the tests were piloted in 1975 and 1976 on a voluntary basis. Even though the State Board of Education acted in 1977 to expand the assessment program to include a tenth grade assessment, it was not until 1979 that the Legislature funded the program. While the Legislature was originally not convinced of the value of the expanded MEAP, the large percentage of districts volunteering to participate in 1977 and 1978 convinced them to mandate the program in 1979.

Assessment of Other Subject Areas

While mathematics and reading are important basic skills (some would argue the most important skills), schools should and do teach students other subjects. MDE, recognizing this, developed objectives in other areas. Test development has occurred in most of these areas and by now, statewide samples of students have been tested in these areas. The original plans called for the assessment of two subject areas each year (in

addition to mathematics and reading) at grades four, seven and ten through statewide sampling to produce an overall picture of the state. Assessment in each area then would follow a four-year cycle continuing to assess all subject areas.

Forces For Change

The MEAP has continued from 1979 to 1985 to assess all fourth, seventh and tenth graders annually in mathematics and reading. In addition, one or two subject areas were selected for sample testing each year. While achievement has risen in mathematics and reading, there have not been appreciable changes in student performance in the areas where only samples of students were tested. Considerable support was evident for MEAP and for changing the program to support instructional improvement in all subject areas tested.

A major force for change of MEAP, of course, has been the spate of reports on the condition of education nationally and in Michigan. A number of these have proposed using testing not only as vehicle to monitor student achievement but also as stimulus for educational reform. In Michigan, for example, a special report written by State Senator Sederburg and Michigan State University Professor Rudman, was prepared that examined changes in performance for various subgroups of students, particularly at the high school level, where comparative data on students in Michigan and the nation is available using college-entrance tests such as the SAT. This report was written in response to A Nation At Risk and the Michigan State Board of Education plan for the future (A Blueprint for Action, 1984), which included recommendations made by the Michigan High School Commission. The following is taken from the summary of the Sederburg and Rudman report:

Over the past few years, state and federal educational policy has targeted the lower achieving student. This targeting of funds and effort has yielded results. However, it is apparent that, at the same time, we may have neglected the better achieving student. In contrast to the prevailing belief, the brightest students have not succeeded regardless of the educational system.

Consequently, we are calling for a shift in educational policy. We must create an educational system that challenges all young people and develops students to the best of their abilities. Emphasis on testing for basic skills for high school graduation and grade promotion reinforce the attitude that teachers and administrators should be most concerned with the lower achieving student. While it is worthwhile to insure that all students possess "essential" skills before graduation, we must not overlook the student who is not challenged by such minimal objectives.

The recent proposals made by the State Board of Education go a long way toward accomplishing the goals outlined here. However, the entire focus must be shifted away from minimal skills which tend to bring high achievers down while trying to bring everyone up to the highest level possible. The State Board and the legislature will need to clarify their philosophical direction as well as set specific goals for whatever educational reform they wish to achieve in the 1980s.

Proposals for Change in MEAP

The Sederburg and Rudman paper contained the first proposals for developing a higher-level test. Although the State Board of Education's report included changes for the assessment program, such changes dealt only with broadening the scope of MEAP to include periodic, every-pupil testing of other subject areas including health, science, career development, and social studies. The State Board of Education has approved the voluntary testing of Health in 1985 and the every-pupil testing of science for 1986.

The Sederburg-Rudman article, however, dealt specifically with higher-level assessment by suggesting, among other things, that:

1. The testing program of the State Board of Education should be changed to adequately measure all Michigan students, not just those below the achievement level determined by the State.
2. The State Board of Education set achievement goals to be attained by all achievement classifications by a specific date. In their "Blueprint for Action" the State Board calls on local boards to initiate a 3-5 year plan to improve achievement. Similarly, the Board should set State goals to improve all categories of Michigan youngsters.
3. State policy should reflect an effort to pressure local school districts to provide programming for the entire spectrum of students. The State testing program should be used to validate or accredit local school diplomas for all students.

- a. *Achievement tests administered as early as the tenth grade should point to areas for potential remediation. The 10th grade test should emphasize reading, language, and basic math skills.*
- b. *An 11th grade exam should include physical science, biological science, and social science. The 12th grade year would be used to assist students who did not meet essential skills in the 10th and 11th grade exams.*
- c. *The State Board of Education should use these tests as the basis for accrediting high school diplomas.*

A response to the Sederburg and Rudman paper by the MDE suggested other possible directions for the MEAP, including expanding the program to periodically assess a third subject area at grades four, seven and ten. In addition, the MDE proposed:

The other way in which MEAP may change in coming years is to assess students beyond the basic skill level. This discussion presumes that (1) testing basic skills is valid and will still be carried out, (2) testing higher-level skills should emphasize the same purposes as the regular MEAP program (i.e., individual student assistance, curricula review and revision, reporting to various audiences), (3) students should be identified based on their basic skill achievement, (4) such higher-level skills are either more difficult subject matter content, critical reasoning skills or higher-level thinking skills (e.g., analysis, synthesis and evaluation from Bloom's Taxonomy), and (5) the students identified can be offered a school program which meets their educational needs, even as schools are helping students who have not as yet achieved the minimums. The presumption is that schools (and the State) can emphasize both "basic" skills and "advanced" skills and not have to choose one over the other (Roeber, 1984).

MEAP staff proposed a plan that included a two-tier approach, with all fourth, seventh, and tenth grade students taking the basic skill level and those that passed, the higher-level examination. It was proposed that advanced tests be developed at three levels (grades 4-6, given in seventh grade; grades 7-9, given in tenth grade; and grades 10-12, given in grades 10, 11, and 12). Staff also developed a list of technical and policy issues for testing beyond the basic skills.

The Department plan was presented to the State Board of Education in early 1985. After considerable discussion, the State Board approved the MEAP staff plan that a study group be convened to examine issues and to develop a tentative assessment plan.

Developing the Plan for the New Assessment Program

Since late 1984, Department staff have been meeting with a planning group consisting of local and intermediate district educators, college and university specialists and others. Represented on the group are gifted educators, assessment and curriculum specialists, content area specialists (e.g., science, reading), and administrators.

The group has spent a considerable amount of time discussing methods to address student needs, particularly those of students who already pass the current basic skills tests. Very early in these discussions it was apparent that there were sharp differences of opinion regarding the direction MEAP should take. Some members of the advisory group, for example, proposed toughening the current content standards tested in MEAP. Others suggested that tests of critical thinking, critical reasoning, or thinking skills be used.

The group pursued both options. Discussions have focused on what “tougher” standards really mean, how higher-order thinking could be tested and how this program could mesh with the current basic skills program. Others have been examining various approaches to teaching thinking skills, looking particularly at how thinking skills are defined and the implications for testing. While viewed originally as an alternative to the current basic skill program (or, at least, a more difficult extension of it), thinking skills are now viewed as a logical complement to the current program, plus any new program which might be developed.

Recommendation for Change

The planning group agreed that there is a need to assess subject content from a conceptual point of view and to include a broader range of subject matter content. In order to encourage the development of students’ thinking skills, the committee also felt that thinking skills should be assessed within each subject content area. Also, the group felt that MEAP should be broadened to include an every-pupil writing assessment, and

subjects other than mathematics and reading should be assessed each year rather than on the current cyclic program. Taken as a whole, the group recommended:

- 1. Basic skills assessment — continuation and revision of the every-pupil essential skills assessments at grades 4, 7, and 10 in reading and mathematics. The revisions should include the assessment of thinking skills, a broader range of (i.e., algebra in ninth grade mathematics test) and the focus on understanding the concept as opposed to a “right answer.”*
- 2. An every-pupil writing assessment be given;*
- 3. Health, science, social studies, and career development be assessed on an every-pupil matrix-sampling basis. It is recommended (2 and 3) be implemented in grades 5, 8, and 11.*
- 4. Thinking skills should be assessed in all content areas.*

The planning group’s recommendations will be presented to the State Board of Education in early 1986. If action was favorable, it would take years to develop the needed testing materials. It would also take time to prepare local districts to test several subject areas at grade levels not previously assessed. Most importantly, staff would need to define higher order thinking skills, both in general terms and also for each subject area in which it will be tested.

Counterforces Against Change

Following the completion of the planning group’s work, the recommendations were presented to the State Board of Education in March, 1986. They received the planning group’s report and referred it to the State Board of Education-appointed advisory council for the service area of the Department in which MEAP is located. This advisory council — the Office of Technical Assistance and Evaluation (OTAE) Advisory Council — is comprised of official representatives of major professional groups such as teachers, principals, administrators, school boards, curriculum groups, as well as technical specialists. The purpose of the OTAE Advisory Council is to advise staff and the State Board of Education on the major issues facing the Office.

The OTAE Advisory Council reviewed the planning group's recommendations and, in May, 1986, voted to oppose the plan and, instead, support a plan that would call for MEAP to develop item banks which local districts could use, in addition to available tests and MEAP tests in the five areas covered by the plan to test one or more of them on a voluntary basis. MEAP would develop, with the assistance of technical groups, standards for equivalence among the various measures used in any subject area. However, testing would not be mandatory.

During the summer, MEAP staff convened an ad hoc group comprised of a subset of the planning group and the OTAE Advisory Council to attempt to develop a compromise which all groups could support. The planning group's recommendations were particularly opposed by four groups: the Michigan Education Association and the Michigan Association of School Boards, both of which feared loss of control of schools, the Michigan Association for Supervision and Curriculum Development, which felt testing was not the proper vehicle for curriculum change and the Middle Cities Association, which felt that state testing duplicated local testing and that the latter was preferable. These groups and others were asked to serve on the ad hoc group.

The group met four times during the summer of 1986 and held several stormy sessions to arrive at the compromise. This compromise was that local districts would be required to give the expanded testing at grades 5, 8 and 11 in writing, health, science, social studies and career development once every four years (but volunteer on off-years) and financial incentives would be sought for participating schools to use for school improvement activities.

During the fall, 1986, the compromise plan was re-submitted to the OTAE Advisory Council, with the interest of sending it to the State Board of Education. Each Advisory Council member was asked to discuss the compromise plan with the organization they represented. In October, 1986, the Advisory Council took formal action on the compromised plan and rejected it. Most major organizations continued to oppose it, even

though the representatives that had served on the ad hoc group had (personally) agreed to the compromise. "Mandatory" testing was the key to the rejection of the compromise.

Final Plan for the Future Approval

Following the vote of the Advisory Council, MEAP staff were informed by the State Superintendent that, with the opposition of about all groups to mandated expansion, he would not put any plan mandating expansion before the State Board of Education. MEAP staff then rewrote the plan for the future to delete any mandated expansion. Instead, the plan calls for the development of tests in health, science, career development and social studies, grades 4, 7 and 10, which are to be offered annually on a voluntary, state-paid basis to local districts. In addition, a writing test will be developed for grades 5, 8 and 11 and offered on the same basis. Staff will continue to develop a program of financial incentive to encourage schools to give the tests and to use the information to review curricula and improve instruction.

This plan was presented to the State Board of Education in March, 1987, and approved unanimously. Tests in the areas of health, science and career development will be offered to districts in the fall, 1987 MEAP; tests in social studies and writing are in development and will be added when ready.

Summary

The MEAP has been in operation since 1969. During that time, it has shifted from a norm-referenced to an objective-referenced program. While the program was controversial in its early years, the emphasis on providing data helpful to improving student learning has helped to improve the support for the program. Grade 10 assessment was added in 1979 to the original grade 4 and 7 programs. In more recent years, periodic, every-pupil tests in other areas, such as science, were proposed. The first area of such testing is science scheduled for 1986.

The recent reports on education have led to a number of suggestions for changing MEAP. These include toughening the basic skills tests, adding measures of critical thinking, and increasing the number of subject areas tested. Staff plans to implement these ideas were presented to the State Board of Education in 1986 and referred to the State Board of Education appointed Advisory Council. The plans were rejected by the Advisory Council. A compromise plan, which contained an element of mandatory testing, was also rejected by the Advisory Council. Consequently, a plan to expand MEAP on a voluntary-state-paid basis was proposed by staff and approved by the State Board of Education. The plan will be implemented beginning in the fall of 1987.

REFERENCES

- Roeber, Edward D. "Michigan Educational Assessment Program: Proposed Plans for the Future" (Michigan Department of Education, 1984).
- Sederburg, William A. and Herbert C. Rudman. "Educational Reform and Declining Test Scores," Michigan School Board Journal, vol. 30, No. 24, April 1984 pp. 8-10.
- State Board of Education Better Education for Michigan Citizens: A Blueprint for Action, (Lansing MI, 1984).