

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.



University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9129451

**The perceived effects of the Michigan Educational Assessment
Program on small, rural school districts in the Upper Peninsula
of Michigan**

DeMarte, Faye, Ph.D.

Michigan State University, 1991

Copyright ©1991 by DeMarte, Faye. All rights reserved.

U·M·I

**300 N. Zeeb Rd.
Ann Arbor, MI 48106**

THE PERCEIVED EFFECTS OF THE MICHIGAN EDUCATIONAL ASSESSMENT
PROGRAM ON SMALL, RURAL SCHOOL DISTRICTS IN THE
UPPER PENINSULA OF MICHIGAN

By

Faye DeMarte

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Educational Administration

1991

ABSTRACT

THE PERCEIVED EFFECTS OF THE MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM ON SMALL, RURAL SCHOOL DISTRICTS IN THE UPPER PENINSULA OF MICHIGAN

By

Faye DeMarte

The researcher's purpose in this study was to investigate the effect of the Michigan Educational Assessment Program (MEAP) on the curriculum in small, rural school districts. Third-, fourth-, sixth-, seventh-, ninth-, and tenth-grade reading and math teachers in ten selected districts were asked to respond to a 33-item forced-choice survey in spring 1990. Personnel in four of the ten survey districts were interviewed to substantiate and enrich data collected with the survey. The data were analyzed to provide answers to two research questions, to indicate the levels of use of the MEAP, and to provide a description of small, rural schools.

Given the limitations of the study, the following major conclusions were drawn: Districts used MEAP results "some" to determine the strengths and weaknesses of math and reading curricula; principals provided "some" leadership in the interpretation of the results; results were communicated "quite a bit" to students, parents, and school boards; and results had

Faye DeMarte

altered course content, teaching methods, and preparation for tests. Districts varied significantly on one-third of the survey items. A number of practices were used in small, rural school districts to raise students' MEAP scores. The four vignettes constructed from the interviews revealed that although these districts had much in common with respect to isolation, student enrollment, type of community, and academic history, the districts varied significantly on curricular issues.

Dissertation Director: Dr. Samuel A. Moore II

Copyright by
FAYE DeMARTE
1991

ACKNOWLEDGMENTS

I express my gratitude to the many people who lent their support and encouragement in completing the requirements for the doctoral program, especially:

Dr. Samuel Moore II, my advisor and dissertation chairman, who provided direction and meaning throughout this academic endeavor.

Drs. Blackman, Groty, and Rowan, dissertation committee members, who focused and refined this study.

Small, rural school district educators who completed the survey and allowed interviews; without them, this study would not have been possible.

DeTour Area School board members and educators, especially Superintendent Jay VanHoven, who granted the time and personal support to see this dissertation to completion.

My family, Bob, Roberto, and Katrina, who sacrificed family time and resources for this pursuit. Their sacrifices, together with their faith and confidence, make this achievement a mutual accomplishment.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
 Chapter	
I. STATEMENT OF THE PROBLEM	1
Introduction	1
Rural Schools	3
Purpose of the Study	8
Importance of the Study	8
Methodology	9
Delimitations	10
Definition of Terms	11
Summary and Overview	12
II. REVIEW OF LITERATURE	14
Introduction	14
Measurement and Evaluation	15
The Popularity of Assessment	18
Tests and Their Use	20
The Curriculum and Testing Link	23
Research on Testing	26
Research on Minimum-Competency Tests	29
The Evans Study	30
The Cruse Study	30
The Rankin Study	31
The Sandifer Study	31
The William Study	32
The Smedley Study	32
The Walstad Study	33
General Findings	33
Testing in Michigan	34
MEAP Studies	37
Early MEAP Studies	40
Dissertations on MEAP	41
Summary	46
Chapter Summary	47

	Page
III. METHODS AND PROCEDURES	49
The Population and Sample	49
The Survey Instrument	52
Collection of the Data	53
Data Analysis	56
IV. PRESENTATION OF DATA	58
Introduction	58
Survey Returns	59
Interviews	59
Results of Data Analysis for Test Usage	60
Questionnaire Item 1	60
Questionnaire Item 2	65
Questionnaire Item 3	69
Questionnaire Item 4	71
Questionnaire Items 4b, 4c, and 4d	75
Use of the 1989 MEAP Results	77
Comparison With Jencka Study	84
Summary of the Use of MEAP Results	86
Questionnaire Item 6	87
Summary of Communication of MEAP Results	95
Questionnaire Item 7	96
Questionnaire Item 8	105
Four Vignettes	107
Aholahiti	109
Koskitown	114
Makivil	118
Vilmi	121
Summary	125
Chapter Summary	125
V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	128
Summary	128
Purpose of the Research	131
Methodology	132
Findings	132
Conclusions	138
Recommendations	141
Further Study	141
Process Guidelines	142
APPENDICES	
A. SURVEY	144
B. LETTER OF PERMISSION TO CONDUCT STUDY	149

	Page
C. COVER LETTER	150
D. TABLES	151
BIBLIOGRAPHY	155

LIST OF TABLES

Table	Page
1. Number of Surveys Sent to and Received From Participating Districts	55
2. Responses to Item 1: Standardized Tests Provide Important Information in Assessing Your District's Effectiveness, by Individual Districts .	61
3. Percentages for Each Item Response in Question 1-- Standardized Tests Provide Important Information in Assessing Your District's Effectiveness-- Across All Surveyed Districts	64
4. The Use of MEAP Results With Other Tests, MEAP Alone, or Not At All by the Individual Districts .	66
5. The Use of MEAP With Other Tests, MEAP Alone, or Not At All Across All Districts	69
6. Responses to Item 3: District Selects or Develops Tests for Items Included in the Local Curricu- lum But Not Measured by MEAP	69
7. Percentages for Each Item Response in Item 3-- District Selects or Develops Tests for Items Included in the Local Curriculum But Not Measured by MEAP--Across All Districts	71
8. Responses to Item 4: Curriculum Guides Are Available in Reading and Math	72
9. Percentages for Each Item Response in Item 4-- Curriculum Guides Are Available in Reading and Math--Across All Districts	74
10. Results of Analysis of Variance for Item 5e: To determine Instructional Priorities	80
11. Results of Analysis of Variance for Item 5f: To Determine Placement of Students in "Remedial" Programs	82

	Page
12. Results of Analysis of Variance for Item 5i: To Identify Staff-Development Needs	84
13. Mean Ratings of the Use of MEAP Results for Specified Purposes by Rural Educators and by Principals in Michigan	85
14. Results of Analysis of Variance for Item 6a: The Principal Provides Leadership in the Interpre- tation of MEAP Test Results	88
15. Results of Analysis of Variance for Item 6b: MEAP Has Been Discussed at Teacher Meetings	90
16. Results of Analysis of Variance for Item 6c: MEAP Results Have Been Discussed at School Board Meetings	91
17. Results of Analysis of Variance for Item 6d: Inservice Activities Have Been Provided to Help Incorporate MEAP Objectives Into Instructional Programs	92
18. Results of Analysis of Variance for Item 6e: Our Staff Shares Information Regarding Teaching Strategies Used to Increase Student Test Scores on MEAP	94
19. Results of Analysis of Variance for Item 7b: Changes in Course Content Based on MEAP Results . .	97
20. Results of Analysis of Variance for Item 7g: The Change in the Frequency of Curriculum Revision Based on MEAP Results	103
21. Results of Analysis of Variance for Item 8a: Teachers Are Familiar With MEAP Test Questions . .	105
D.1. Percentages: Curriculum Guides Are Available in Reading and Math	151
D.2. Means for All Variables on Survey Instrument	152
D.3. Means for Significant Contrasts	153
D.4. Statistically Significant Contrasts Within Districts for Questions That Are Statistically Significant Between Districts	154

CHAPTER I

STATEMENT OF THE PROBLEM

Introduction

Before 1970, the word "accountability" rarely appeared in educational publications, nor was it found on agendas for educational organizations. Accountability, as applied to the educational process and outcomes, was pervasive in the 1980s, with small, rural schools striving for greater effectiveness. The response of small, rural schools in the Upper Peninsula of Michigan to the curricular accountability issues, and more specifically, the Michigan Educational Assessment Program, was the focus of this research.

According to Tyler (1971), three recent developments appear to have influenced the accountability movement: the larger portion of the family's income that is spent on taxes, the larger number of teenagers who are unable to meet literacy standards required for employment, and the success of management processes in industry.

Schools were asked to justify increased budgets by showing greater educational effectiveness. Educators were told that they must be held accountable for greater results if taxes were to be increased. More specifically, said Lessinger (1970), the public served notice to the schools that society members expected all

children to learn at least the basic skills, that failures were to be regarded less as the fault of the child or of his/her background than of the school, and that the proper response to failure was not excuses but reform.

With the measurement of basic skill attainment came the development and extended use of standardized tests. Brown v. Board of Education, the Elementary and Secondary Education Act, the Coleman Report, the National Assessment of Educational Progress, and the "back to basics" movement brought forth activity with respect to equity and accountability of results. There was also a trend toward a stronger national and state role in education. All considered assessment programs in some form. In 1978, 33 states initiated minimum-competency measures. Although state actions have not always resulted in minimum-competency tests, states have increased their role in the education of school children. Such a role affects the K-12 curriculum.

In 1970, the Michigan Department of Education began to develop the Michigan Educational Assessment Program (MEAP) "to assess what educational attainments and deficiencies exist" (Porter, 1976, p. 668). MEAP was the selected tool that would measure student, school, and district effectiveness. Adjustments in expenditures, curriculum, and delivery were to follow if districts were found deficient. Thus, according to Porter, Michigan could more assuredly "guarantee that all students, without respect to race, income, or social class, will acquire the minimum school skills necessary to

take full advantage of the choices that accrue upon successful completion of public schooling" (Lessinger, 1971, p. 42).

A decade later, Donovan (1982) of the Department of Education boasted that "state assessment has been accepted as a good indicator of the level of achievement in Michigan and a valuable curriculum tool" (p. 9). House, Rivers, and Stufflebeam (1974) suggested that MEAP is flawed in that this accountability model does not serve local decision making. Local autonomy is crucial to districts and is not easily relinquished to the state.

Rural Schools

Data about small, rural schools "tend to be scanty, and, at worst, simply non-existent" (Sher, 1977, p. 3). In addition, rural people, rural communities, and rural conditions are diverse. The commonality of "an island village off the coast of Maine, a coal mining town in West Virginia, a ranch area in Wyoming, a college town in Minnesota, an impoverished community in the Mississippi delta region, a ski resort in Vermont, a migrant worker settlement in Texas, an Alaskan native village near the Arctic Circle, and a prosperous grain farming area in Iowa" (Sher, 1978, p. 3) is that they are all rural. Rural America is heterogeneous and complex; thus, a simple solution to increase the educational effectiveness of all rural schools seems highly unlikely.

The National School Board Association defined rural schools as those schools with a student enrollment of fewer than 2,500 or located in a rural setting. According to the National Rural

Development Institute, an area is rural when the number of inhabitants is fewer than 150 per square mile (DeYoung, 1987). Jess (1984) defined rural as those who live in the open countryside and communities with fewer than 2,500 inhabitants. He further stated that the commonality of all definitions is that rural schools are small in enrollment size, are a great distance from any major urban center, and are an extension as well as a reflection of the community they serve.

In the studies conducted in rural schools, student enrollment size tended to vary due to the manner in which rural was defined. However, Barker and Muse (1983) conducted a comparative study of districts with enrollments of 300 or fewer students versus districts with enrollments between 301 and 900 students. They found that districts with enrollments of 300 or fewer students experienced greater difficulties attracting personnel, tended to have greater financial difficulties, had fewer support services, had more teachers teaching outside their area of certification, and had fewer curricular offerings. Economies of scale were not engaged in districts with student enrollments under 300.

Currently, in Michigan, legislators have defined "small" as those districts with enrollments of 500 or fewer students. This definition is used for "recapturing" purposes: These districts do not have to return to the state coffers those funds received in the categories of transportation, special education, and so on, even though they are spending more per student than the school financial-aid formula dictates. It is believed that these schools are not

able to take advantages of the economies of scale because of their student enrollment size.

Rural schools vary not only to the extent that they are small, but also according to the type of community in which they are located. Gjeltén (1982) offered the following typology:

Stable rural--Classic rural community--prosperous, peaceful, traditional, and mostly white.

Depressed rural--Economic insecurity abounds, outmigration is high, local economy is often underdeveloped and there is a moderate to high minority population.

High-growth rural--Faces problems inherent in rapid growth: inadequate school facilities, housing, and services.

Reborn rural--Attract a refugee population from the city seeking a rural lifestyle; they are zealous defenders of many traditional rural customs and institutions.

Isolated rural--Have many characteristics of the other types, but isolation leads to separate problems (transportation, commerce, and cultural activities are all affected).

Although they are difficult to define, rural schools are not so elusive as to evade the pressure by their constituents to become more accountable. This pressure compounds the task of education in small schools; they must deal with the factors of ruralism as well as the historic problems associated with small, rural schools. These problems include absenteeism and nonenrollment, recruiting highly competent teachers, providing special education and other specialized services, securing needed capital, and compensating for inherent isolation and sparsity of population (Muse, Barker, & Smith, 1983; Sher, 1977).

Are rural schools institutions of educational excellence? "Achievement on the cognitive measures has demonstrated the most consistent pattern, with the suburban areas scoring highest, followed by rural-small town communities, and last of all urban areas (Kohr, 1983). Superintendents of small, rural schools indicated that 61% of their students' performance was "close to the national average," whereas 27.8% reported performance "above the national average" (Muse et al., 1983). In Michigan, using MEAP as a cognitive measure, small, rural school districts have been found in the top 20 as well as in the bottom 20 ranks (Hornbeck, 1989).

This ranking supports Sher's (1977) declaration that "rural schools can be excellent and be as productive as anywhere in the nation--such excellence will not occur by default, but rather by deliberate actions aimed at building upon what they have, as well as providing for what they lack" (p. 19). Further, Nachtigal (1980) stated that accepting rural reality would not mean that the same educational program offered in the urban school should be provided in the rural community, for the situations are different and so are the needs. "Our best strategy in the small schools arena for both improvement and survival is not to compete in program breadth with larger schools but to compete equally on the issues of quality such as those identified in the school effectiveness research" (Marshall, 1986, p. 7).

Buttram and Carlson (1983) discussed the seven characteristics of an effective school as they pertain to rural schools. The strengths of small, rural schools lie in a safe, orderly

environment; high expectations; opportunity to learn; and home-school relations. Small rural schools do not do as well in focusing on a clear mission and providing instructional leadership. Coupled with frequent monitoring, a clear statement of purpose, and instructional leadership are recognized elements of curriculum development and implementation.

In Barker's (1985) study, superintendents ranked the major challenges facing them in their small, rural districts. Improving the curriculum was a close second to obtaining adequate financial support. Curriculum improvement "can best be furthered through the encouragement and support of networking and technical assistance" (Gjelten, 1979). Often encouraged are long-distance learning and shared services in vocational education, special education, counseling, and curriculum development (Loustaunau, 1975). The curriculum-textbook-test match seems not to have received attention in literature on rural schools.

Although the effective schools research has emphasized an administrative function to curriculum development and implementation, McClurkin (1970) reminded the educational community that the establishment of minimum standards depends more on the classroom teacher than any other factor. The factors influencing teacher instructional changes have been investigated (Floden, Porter, Schmidt, Freeman, & Schwille, 1981), and the presence of objectives was found to be the strongest force to cause change in course content.

This study was undertaken to provide information as to the status of MEAP, the mandated state assessment test, in small, rural school districts in the Upper Peninsula of Michigan. District practices in curriculum development and test alignment were also investigated. In addition, the researcher attempted to answer the question: Do small, rural districts teach to the test, or do they create a curriculum designed to meet their local needs?

Purpose of the Study

The researcher's purpose in this study was to investigate the effect of MEAP on the curriculum in small, rural school districts. The effect was assessed by investigating the testing program, the curriculum, classroom use of test results, and the communication of test results. Subsumed under the main question are several subquestions:

1. Are there significant differences among districts with respect to the use of MEAP results?
2. If MEAP results are used, how are they used?
3. What practices, if any, are intentionally used by small, rural schools to respond to the MEAP?

Importance of the Study

One of the stated purposes of the MEAP is to improve student achievement. This should occur and be evident if local school districts use assessment results to revise curricula and improve instruction. On a statewide basis, Aquino (1975), Bushaw (1988), Jencka (199), and Steele (1976) found that MEAP was used in

determining the general level of students' achievement, as well as the strengths and weaknesses of the curriculum. MEAP was used to determine which "essential" skills were taught, to report to the board of education, or to use in conjunction with other tests.

In only Jencka's (1990) study were data analyzed across strata. The sixth stratum was considered rural, encompassing the northern Lower Peninsula and the Upper Peninsula. As a stratum, the schools in the sixth stratum did not differ significantly from other strata. However, this stratum includes schools with student enrollments of 1 to 4,840. Is variability in MEAP use present with such variability in enrollments?

In addition, are State Department of Education mandates able to predicate change in small, rural districts isolated geographically from other districts and from the State Department? If school districts are affected by MEAP, what has changed as a result of this mandate? Do any districts intentionally try to raise student achievement scores?

This exploratory research was designed to examine these questions and to provide educators with some information as to the status of MEAP in the small, rural school districts in the Upper Peninsula of Michigan.

Methodology

To determine the use of MEAP in small, rural schools, a random sample of ten school districts with enrollments of 500 or fewer students was selected. The population of such districts was

recorded in the 1990 Michigan Education Directory and Buyer's Guide. Twenty-one such K-12 districts exist in the Upper Peninsula of Michigan.

In the ten selected districts, superintendents, principals, and third-, fourth-, sixth-, seventh-, ninth-, and tenth-grade teachers received a survey containing 33 forced-choice items. One question, consisting of ten items, was replicated from the Jencka (1990) study.

Data were also collected through extensive interviews in four of the ten selected districts. Interviews were conducted to validate and extend the findings of the survey. Interviews were not limited to those individuals who received the survey.

The collective responses of district personnel were the unit of analysis. This analysis was used (a) to determine whether there were significant differences among districts with regard to MEAP use, (b) to determine how MEAP was used, and (c) to identify what, if any, practices small, rural schools were using to raise students' achievement scores.

Delimitations

Delimitations of the study were as follows:

1. Data collected were based on small, rural Upper Peninsula school districts that had a total enrollment of 500 or fewer students. This enrollment figure is currently used by the State of Michigan for districts to be exempt from recapture of funds due to the districts' inability to participate in economies of scale.

Recapture is the payment by the local school district to the state of those categorical funds (transportation, special education, and so on) that the district receives because the district is spending more per student than the state-aid formula dictates.

2. The data for the study were based on the perceptions of superintendents, principals, and third-, fourth-, sixth-, seventh-, ninth-, and tenth-grade teachers rather than on observed behavior or documents.

3. The teachers' perceptions were limited to those who taught reading and math.

Definition of Terms

Curriculum. The body of knowledge that is presented in formal statements, as well as that which is presented in actual educational experiences.

Effect. An effect includes the intentional as well as the unintentional results of MEAP scores on teaching and the curriculum.

Michigan Educational Assessment Program. The program developed and administered by Research, Evaluation and Assessment Services: Michigan Department of Education. It is defined in Act 307 of the Public Acts of 1969 and by Act 38 of the Public Acts of 1970. The MEAP tests are administered to fourth-, seventh-, and tenth-grade students in the academic areas of reading and mathematics. Fifth-, seventh-, and eleventh-grade students are given MEAP tests in science. The tests are analyzed and returned to the local school district for use.

Utility. Utility implies worth. MEAP results possess utility if they are useful in evaluating what is taught, and if they are a component of measuring student outcomes in an effective school plan.

Summary and Overview

The need for research in small, rural schools was established in this chapter. A brief discussion of accountability issues and the MEAP was included. Rural was defined, and a community typology was given. Rural school problems and challenges were discussed. Also included in this chapter were a statement of the problem and the purpose for the research. Definitions of terms and the methodology employed in the research were presented.

Chapter II contains a review of literature relevant to the study: equality and accountability, the popularity of assessment, tests and their use, curriculum and testing link, research on testing, minimum-competency testing, testing in Michigan, and MEAP studies.

Chapter III contains a description of the population and the sample, a description of the survey instrument, data-collection procedures, interview questions, and data-analysis procedures used in the study.

Chapter IV contains a review of the findings; the general level of use is reported as a mean, significant findings are discussed, and a portrait of four small, rural schools is provided.

Chapter V contains a summary of the study, followed by conclusions. Recommendations for further research are also provided.

CHAPTER II

REVIEW OF LITERATURE

Introduction

A number of sweeping changes have occurred in education during the last 25 years in response to public demands. In the 1960s and 1970s, public dissatisfaction was widespread with regard to social promotion, graduation based on attendance rather than achievement, and a curriculum that was perceived to be watered down. Declining test scores, grade inflation, complaints about the skill levels of high school graduates, and the rising costs of public education compounded this dissatisfaction (Pedulla & Reidy, 1979). The proper response to these failures was reform, not excuses (Lessinger, 1970).

Changes did occur, and they included (a) a greater role in local K-12 educational programs by national- and state-level policy makers, (b) an emphasis on equitable treatment of students, (c) a focus on learning and teaching processes, and (d) a greater concern for accountability for what children should learn. The role of testing, particularly state assessment tests, has increased as a result of these changes (Airasian, Madaus, & Pedula, 1979; Bushaw, 1988).

Measurement and Evaluation

Measuring and evaluating the progress of students is not new (Mehrens, 1979). In 1856, an examination was used for admission into high school and for promotions. In 1897, Joseph Rice began his spelling tests. The shift from competence to educational efficiency and achievement occurred in 1918. According to Jaeger and Tittle (1980), the 1930s ushered in the new concept of educational evaluation, supplanting the testing of individual students. After World War II, the demands on the public school system increased because of a larger student enrollment and expanded opportunities (Jaeger & Tittle, 1980). Resnick (1985) concluded that "tests and examinations have traditionally served as a major means of setting and maintaining educational standards, and they do this by monitoring the performance of both educational institutions and the individual students in them" (p. 178). He further stated that "there is abundant evidence that the overriding role of educational testing has been to serve purposes of public accountability, program evaluation, and institutional comparison" (p. 11).

The educational tenor of the 1960s and 1970s was that of equality as well as equity. Brown v. Board of Education (1954) addressed equal educational opportunity as measured not only by inputs but also by outcomes. San Antonio v. Rodriguez (1971) addressed financial inputs and their relation to student outcomes. Both cases emphasized the measurement of student outcomes, causing educators to search for ways to measure student skills. The Elementary and Secondary Education Act (1965) heightened this search

for student outcome measures by providing financial assistance to low-income, low-achieving schools with the requirement that these programs engage in systematic evaluation (Wise, 1979). Thus, these funds provided greater educational opportunity while expecting success to be monitored and measured by tests.

In 1966, the Equality of Educational Opportunity report (Coleman et al., 1966) was released by Coleman and his colleagues. One conclusion of the report was that school factors had little effect on students' standardized achievement test scores. Differences in achievement resulted from factors outside the school, such as the home, the neighborhood, and peers. The report focused on outcomes, as measured by standardized achievement tests, and set a precedent for a new type of research that would focus on what the student leaves the educational process with, rather than what is put into the educational process.

Reaction to the Coleman Report caused a series of studies to be undertaken because groups in the educational community believed that processing variables made a difference. The National Assessment of Educational Progress, developed in 1969 by the Education Commission of the States, was an instrument "created to examine achievement in 10 learning areas, to spot changes in levels of achievement, and to apply the implications of those changes to national educational policy" (Wise, 1979, p. 9). Numerous studies followed, some of which became part of the effective schools research.

Curricular content, that of the basics, was demanded in the 1960s and 1970s because of reports of

. . . declines in SAT scores, of unfavorable comparison of achievement levels among American youngsters and their counterparts in Japan and European industrial nations, of the lack of teachers adequately prepared to teach math and science, . . . of high school graduates who are functional illiterates, all have been producing a rising sense of alarm among the American public. (Hall, Jaeger, Kearney, & Wiley, 1987, p. 16)

The Gallup polls of the 1960s and 1970s clearly indicated that the public wanted the basics: reading, writing, and arithmetic (Wise, 1979; Womer, 1981). The back-to-basics movement had one primary goal, that of producing pupils who were well grounded in basic skills. Secondly, it sought to motivate students and teachers to work harder, to increase the cost effectiveness of education, and to instill school accountability (Airasian, 1979). Lessinger (1970) stated that "the public is demanding that school officials guarantee the acquisition of basic skills and account for public money in terms of certified educational results."

The court cases involving educational equity, the Elementary and Secondary Education Act, the Coleman Report, the National Assessment of Educational Progress, and the back-to-basics movement shifted the focus of educational evaluation from equity of input to equality of output. These outputs, or educational outcomes, historically have been measured by tests. This new focus, or "this education reform effort . . . mandated a new form of testing, or expanded uses of existing testing" (Haney, 1985, p. 4). This reform effort brought forth an increase in student testing and the development of criterion-referenced and minimum-competency tests.

The Popularity of Assessment

Educational policies of the 1960s and 1970s share a common rationalistic set of assumptions about schooling. These assumptions are that (a) a limited set of measurable goals for education can be determined, (b) tests can be designed to assess the performance of these goals, (c) students of various abilities can be educated efficiently, and (d) rules and procedures are superior to the exercise of judgment by local educators (Wise, 1979). The goals of education, as in the first assumption, were initially defined as the basics of reading, writing, and arithmetic in response to the public demands of that time.

States responded, taking their role in education more and more seriously. The luxury of leaving curriculum to the education experts could no longer be tolerated as states experienced the "loss of basic industries to allegedly better educated foreigners" (Kirst, 1987, p. 9). In the years between 1963 and 1974, state legislators enacted at least 73 laws (Popham, 1985). Their concern seemed to be that of ensuring adequacy in educational achievement rather than quality in educational opportunities (Wise, 1979). The strong American tradition of local autonomy was compromised, with states doing more monitoring and less consulting.

Tests were used for two obvious reasons. According to Lazarus (1981), a test is easy to defend politically: "The state can take the position that it merely wants to monitor student progress in the basic skills, leaving the means for doing so to the local

authorities" (p. 11). Second, a test is relatively inexpensive, conspicuous, and instituted quickly.

In addition to being the least intrusive means of gaining control from afar, tests were selected for a number of other reasons. Tests had traditionally served the functions of selection, placement, and outcome comparisons (Haney, 1985). The term "achievement testing" was understood; people made a quick association to the Scholastic Aptitude Test, the Iowa Test of Basic Skills, or the California Achievement Tests. In fact, the public would probably define "assessment" as the "measurement of students' achievement by standard multiple choice test items for the purpose of showing the public whether or not schools are doing the job" (Stiggins, 1985, p. 69).

In 1977 the National Association of School Boards (NASB) surveyed its membership regarding testing. The respondents indicated that 75% of their districts had districtwide testing programs; these programs were more prevalent in suburban and urban districts or in districts with student enrollments of 1,000 or more. In addition, one-third of those responding indicated that every student was tested, with very few districts practicing random selection. Porter (1983) found that 93% of the nation's school districts used standardized tests of mathematics in the elementary grades. Standardized tests seemed "normal" and went unexamined.

Accountability, using test scores, ranged from a complete analysis of a system's programming, to simple reporting of student performance, to no reporting. However, all districts participating

in compensatory education testing were engaged in mandated federal government educational program audits (Smedley, 1987).

Resnick (1985) summed up testing in the following manner: "American children are the most tested in the world and the least examined. . . . They take one or more tests practically every year" (p. 17). Students are tested, but the results would seem to have little effect on classroom instruction.

Tests and Their Use

When state governments mandated testing, they did so without infringing, at least directly, on the local responsibility for curriculum (Airasian, 1979). Their intention was to assess, but assessment took on at least two different meanings.

Teachers inside the classroom define assessment quite differently from those outside the classroom. Teachers use an array of assessments--some standardized and some individualized, some based on paper-and-pencil tests and some based on observation, and some formal whereas others are informal (Stiggins, 1985). In fact, teachers depend heavily on their own observations and judgments, more so than the paper-and-pencil tests (Kellaghan, Madaus, & Airasian, 1982; Salmon-Cox, 1981). These classroom assessments influence classroom learning (Stiggins, 1985). Standardized test results only confirm the teacher's perception of the pupils' scholastic achievement (Kelleghan, 1982; Salmon-Cox, 1981). In a 1979 Teacher Opinion Survey, 77% of the teachers thought there would be no effect on teaching if school systems eliminated the use of all standardized tests, and 62% believed that the most important aspects

of student progress are not measured by tests (Mika, 1982). Salmon-Cox (1981) found similar results; teachers rarely used test information to mold their instruction or curriculum content.

Darling-Hammond and Wise (1985) reported their data on 43 randomly selected, mid-Atlantic teachers in a slightly different manner. Ninety-five percent of these teachers indicated that the use of tests altered their curriculum emphasis. Teachers found it difficult to adapt uniform educational approaches to the perceived needs of students. Although teachers acknowledged the usefulness of a common direction, they also perceived specific, detailed standards as detrimental to individual students.

Why do the standardized tests that mean so little to the classroom teacher command so much attention from legislators and the general public? Several reasons have been proposed, which include differences in informational needs and differing perceptions of the ability of tests to measure "success."

Educational administrators, legislators, and the public have different information needs. They must show or need to know whether the school is doing the job it is supposed to be doing, using consistent, scientific test scores (Stiggins, 1985). Teachers need feedback on classroom teaching.

"Success" to the public is mastery of basic skills, whereas "success" to teachers means growth in the cognitive as well as the social domain (Salmon-Cox, 1981). The practice of testing is further questioned when the mastery of basic skills in math and

reading apparently has little effect on the employment or wages of students entering the labor force after high school (Ekland, cited in Jaeger, 1980).

The use of test results by administrators and the public was confirmed in the NASB survey (1977); districts indicated that they tested for many reasons, but the primary reason was to facilitate decision making. Of those responding, 38% said the board used the scores for "informational purposes only"; 66% used test scores to judge programs, administrators, and teachers; 55% used scores for judging the effectiveness of districtwide programs; and 50% used test scores to help determine curricular changes.

Both the American Federation of Teachers (AFT) and the National Education Association (NEA) have taken a position on how tests should be used. Speaking for the AFT, Gould (1980) indicated:

The proper use of standardized tests is for diagnostic, prescriptive, selection, placement, and program evaluation purposes and test information should not be the sole basis for decision making, but must be analyzed together with a number of other factors.

According to the NEA (1979), standardized testing should be eliminated because of its misuse, which includes the media's "invidious public comparisons of student achievement test scores" and because they might be "used to evaluate teachers."

Clearly, tests have been selected to measure the effectiveness of schools and subsequent comparisons despite teacher dissatisfaction with such use. Teachers prefer those tests that are diagnostic or those that are closely aligned with classroom instruction

(Salmon-Cox, 1981). These teachers resent schools being "graded" by improvement of student scores in the cognitive domain only.

The Curriculum and Testing Link

Remarkably little attention has been given to how the competencies or skills measured on standardized tests relate to the present school curriculum or, more precisely, the match between what is tested and what is taught (Airasian, Haney, & Madaus, 1979). "An important characteristic of the U.S. tests," wrote Resnick (1981), "is that they are typically divorced from the curriculum" (p. 179).

English (1978) reiterated Resnick in saying that "standardized tests do not correlate with the real curriculum or if they do the system is usually ignorant about the degree of correlation" (p. 46). Further, standardized tests are not measures of any specific curriculum; they are composed of "random events which will result in a bell shaped curve. Curriculum comes into being for the exact opposite rationale--a curriculum is a statement of non-random events" (English, 1987, p. 35). Jenkins and Pany (cited in Idol, Nevin, & Paolacci-Whitcomb, 1986) took a more extreme position, indicating that standardized tests "sometimes serve as poor predictors of how well a student might perform in the classroom."

When curriculum guidelines are vague, the selection of tests based on the content of the curriculum is often impossible. A school system that has vague goals and vague expected outcomes has few criteria to judge the effectiveness of which approach best reaches its goals (English, 1978). On the other hand, Popham (1977)

predicted that where certification tests are described with sufficient clarity to permit targeted instruction, the curriculum will become more and more test-valid. Airasian (1979) agreed and indicated that where there is a choice between emphasizing tested and nontested objectives, in general the objective actually tested assumes primacy. "Faced with a choice between one set of objectives which is explicit in the course outline and a different set which is explicit in the certifying examinations, students and teachers generally focus upon the latter" (Airasian, 1979, p. 129).

However, tests are not generally connected with the curriculum taught (Schalock, Felding, Schalock, Erickson, & Brott, 1985), and "the procedures for scoring, managing, and reporting do not address the information needs of teachers and administrators. Further, schools lack policies and procedures to guide inquiry and discussion about the implications of test information (Haney, 1984, p. 640). Haney concluded that school people are not knowledgeable about the technology of testing. Misinterpretations are made on grade-equivalent scores and even percentile ranks. In addition, and fortunately, consequences are not attached to test results.

Small, rural school districts tend to be at a greater disadvantage than urban or suburban schools as they do not have part-time employees to work at evaluation. The task is generally shared among teachers and principals, neither of whom has expertise in evaluation but who must determine a planned response to the results of an evaluation (Lazarus, 1982; McClurkin, 1970).

It must be noted that the extent to which a teacher responds to any urge to increase the overlap between instructional content and test coverage is by and large determined by the extent to which high test scores are a perceived good (LaMahlew & Leinhardt, 1985; Popham, 1978). Many factors influence content selection; all are necessarily filtered through the teacher (Darling-Hammond & Wise, 1985; LaMahlew & Leinhardt, 1985). When textbooks and other instructional materials are considered for selection because of state-sanctioned tests for some specified bit of information, there is precious little hope for autonomy or for genuine improvement in schools (Orlich & Ratcliff, 1977).

Those schools which use tests that are consistent with their instructional focus will appear effective whereas other schools will not (Porter, 1983). But for many schools, "What is tested is what is taught" remains an unexamined truth. Ebel (cited in Rudman, 1977) summed up the testing-curriculum link in the following manner:

Just as there is no warrant for giving away answers to particular questions, so there is no warrant for testing pupil performance on tasks they were not taught to perform. . . . An assessment test must be thoroughly relevant to the instruction it is intended to assess.

Debra P. v. Turlington (1984) affirmed, through court mandate, the need to instruct students on the tested objectives, the objectives in this case of the Florida minimum-competency test. Madaus et al.'s (1979) research in 50 Irish schools further indicated that criterion-referenced tests, the tests that match curriculum to the test, were much more reliable in measuring school effectiveness than were norm-referenced tests.

Research on Testing

In this section, the review of research on testing includes an examination of studies using standardized tests as well as minimum-competency tests.

According to Mehrens and Phillips (1986), "test publishers go to great lengths to define their tests' content so that they have as much in common with as large a set of district objectives as possible" (p. 185). But they further stated that this does not mean that "different districts, different buildings in the districts, or teachers in a school building, will be offering curricula and instruction that match the test equally well" (p. 185). Cooley and Leinhart (1980) indicated that the degree of mismatch can have a fairly large effect on norm-referenced tests. However, the Mehrens and Phillips study indicated that norm-referenced tests detected overall differences in the quality of instruction but did not detect differences at the level of individual objectives.

In 1987, Mathison completed three case studies analyzing the effect of standardized testing on teaching and curriculum in three midwestern school districts. Data were collected from written documentation as well as interviews. Senate Bill 730 in this Midwest state required that all pupils in grades 3, 6, 8, and 10 be tested in reading, mathematics, and language arts. In addition, the tests had to be content valid; that is, the test had to match the locally chosen learner objectives. Mathison found that standardized tests were used primarily for identifying children who might have

learning difficulties. Data were not perceived to be related to curriculum decisions. The exception existed in Middleton, a rural school district with a student enrollment of approximately 5,000. Here the Science Research Associates (SRA) test was used extensively for curriculum review and had a powerful effect on what the students were taught.

Despite the fact that two of the three districts did not perceive test results to be related to curriculum decisions, subtle effects on teaching and curriculum occurred. These effects were as follows: attempts to match material and the test, alignment of classroom question format to test format, and an increased awareness by the teacher of the test content, causing a shift of content from one grade level to another. The teachers felt pressured to raise test scores, which increased their awareness of test content. However, poor test results provided little direction, with many teachers resorting to a review of key concepts before test taking. Some teachers gave additional instruction to brighter students in hopes of raising the overall classroom performance. These results were similar to those found by Salmon-Cox (1981).

The Institute for Research on Teaching (IRT) has conducted a number of studies regarding testing and curriculum. Porter and others (1983) compared fourth-grade math textbooks to standardized tests. They concluded that the tests and textbooks were not always in agreement on what students should achieve.

In another study conducted by the IRT (Porter, 1981), researchers sought to determine which factors caused teachers to

change instructional content. The teachers in this study responded to six vignettes containing the following persuasive factors: district-mandated textbooks, published objectives, published standardized test results, requests from principals, teachers from higher grades, and requests from parents. The strongest factor that caused teachers to include or emphasize material was district objectives; the next strongest factor was published test results. These research findings were similar to Walstad's (1984) findings.

Cohen (1987) reviewed four studies investigating instruction and test alignment. The Koczor, Tallarico, Fahey, and Elia studies yielded the following conclusions: Instructional alignment caused a 4:1 effect, what to teach was paramount to how to teach, and poor test scores had more to do with misalignment of what teachers taught and what they thought they taught (Cohen, 1987).

Edelman (1981) sent questionnaires to 130 third-grade teachers from 39 schools in the Los Angeles area. These teachers were asked to respond to questions about the California Assessment Program, the Comprehensive Test of Basic Skills, the Survey of Essential Skills, and the Basic Inventory of Natural Language. Teachers thought they were moderately prepared for these tests, with percentages ranging from 62% to 74%. Thirty percent of the teachers taught the specific information included on the test, whereas 60% of the teachers emphasized the subject matter throughout the year. Although knowledge of test content did not alter teaching methods, 20% of the teachers spent more time with the slower students. The majority of

the teachers indicated that the testing program was not an important part of the school program and that the information gathered could have been obtained using fewer tests and in a more efficient manner--namely, through teacher observations.

In summary, the research on testing indicated that although teachers did not see test data directly affecting curricular decisions, they altered their behavior because of testing in their district. The most effective means to get teachers to alter their behavior was through district objectives and the publishing of test results.

Research on Minimum-Competency Tests

Airasian (1979) defined minimum-competency testing as "a program in which students are tested to determine their mastery of certain skills defined as essential aspects of school learning or essential for performing tasks routinely confronted in adult life" (p. xvi).

Most of the more than 40 statewide competency programs rely on markedly less stringent sanctions for those who fail than does Florida. The most stringent sanctions would include grade retention or denial of a high school diploma. Less stringent standards include notifying parents or providing remedial assistance. Studies by the following researchers are reviewed in this section: Evans, Cruse, Rankin, Sandifer, William, Smedley, and Walstad.

The Evans Study

Evans (1985) studied the effect of California's mandated minimum-competency test on curricular changes at the secondary level. Senate Bill 813 identified course requirements, specified required standards of achievement in each course, and mandated certain staff-development activities. Evans compared the 1981-82 overall credit requirements for high school graduation and the credit requirements in English to those of 1976-77, the year of the bill's enactment.

Evans found that the minimum-competency test had been a catalyst for change. The overall requirements for graduation remained the same, but there was a return to more traditional programming. In the English course, a greater emphasis was placed on writing. Also, there was an increase in staff-development inservice programs.

The Cruse Study

Four studies conducted by Cruse, Rankin, Sandifer, and William were reported by Popham (1985) as indicators that minimum-competency tests encouraged change and produced higher test results. A brief summary follows.

The Texas legislature passed Senate Bill 350 and, as a part of that law, the Texas Education Agency was directed to adopt and administer criterion-referenced tests designed to assess basic skills in reading, writing, and math of all students in grades 3, 5, and 9. Cruse (cited in Popham, 1985) indicated that in the third

and fifth grades the average improvement was 10%. The ninth grade posted its highest test scores. He credited improvement to increased communication between educators, instruction related to the tested objectives, textbook alignment, and public pressure through the news media. Texas has added the Texas Educational Assessment to its mandated testing legislation, which raises standards, tests more grades, and requires that students pass a twelfth-grade test before receiving a high school diploma.

The Rankin Study

In August 1977, Detroit instituted its Proficiency Program, assessing high school students' basic skills in reading, writing, and mathematics. What resulted, according to Rankin (cited in Popham, 1985), was a modified curriculum, remedial instruction, and different teaching strategies. Not only did the percentage of high school students passing the exam increase each year, but also the students' scores on the mandated Michigan Education Assessment Program (MEAP) test increased. The percentage of high school students passing the exam in 1981 was 74%, and in 1983 it was 79.1%. MEAP scores during the same period went from 29.6% to 38.7% passing at the 75% level in reading and from 51% to 61.5% passing at the 75% level in math.

The Sandifer Study

In 1978 South Carolina established a Basic Skills Assessment Program, setting minimum standards for readiness in grades 1, 2, 3,

6, 8, and 11, as well as administering a criterion-referenced test in reading and math at the end of each of the aforementioned grades. Sandifer (cited in Popham, 1985) reported that, between 1981 and 1984, the increase ranged from 9% to 18%.

The William Study

William (cited in Popham, 1985) reported that the State Board of Education in Maryland launched Project Basic in 1980. Included in this program was the Maryland Functional Testing Program, which mandated the testing of skills in reading, writing, and mathematics. In 1980, 78% passed the reading test; in 1983, 94% passed this test. Similar test results were reported in math, with 40% passing in 1980 and 61% passing in 1983.

The Smedley Study

In October 1976, the State Testing Program of Washington was instituted, which required academic-achievement testing of fourth-grade students in the public schools. In 1985, the program was expanded to include eighth- and tenth-grade students, and the Metropolitan Achievement Test was adopted as the state's official assessment instrument. Respondents to Smedley's (1987) survey indicated that they lacked training and time to use the test results but that the test was a valuable tool to evaluate curriculum and instruction. However, they did not support it as a means of

comparison. They believed the program was not as effective as it ought to be.

The Walstad Study

Walstad (1984) conducted a study to determine whether changes in curricula, testing programs, or teacher training would influence the percentage of students passing a minimum-competency test. He found that the only significant factor that altered test scores was pretesting students before the mandated test. Walstad concluded his investigation by saying, "Curriculum revision and teacher training may be thoroughly justified from the educational perspective, but these indirect methods do not appear to explain differences in MCT performance across districts over time" (p. 266)

General Findings

Floden and others (1988), in investigating districts regarding the test-teaching link, found that administrators were not using all the tools available to them to strengthen this linkage. The single textbook, instructional objectives, and time guidelines were tools that influenced instructional content. Districts reported making heavy use of testing programs and objectives but only moderate use of time guidelines and adoption of a single textbook. They further reported that teachers did not believe they would be rewarded or punished for teaching the prescribed content.

Porter (1989) found that, through keeping logs of what teachers actually taught, "the teacher content practices [were] wholly consistent with the disappointing results of student achievement." The ultimate decision of content was the teacher's (Floden et al., 1981; Porter, 1989). Thus, teachers taught what they wanted to teach, despite mandated tests and district objectives; those teachers who chose to teach test content had students who scored higher than those who chose to ignore such an alignment.

It has been found that students--those who are to benefit from the educational process--simply do not care about test results on minimum-competency tests (Blau, cited in Jaeger, 1980). The poor students, however, saw the tests as an additional barrier to success and self-esteem; the good students saw them as a barrier to efficient use of time.

In summary, it appears that once tests are mandated, students' test scores on these instruments rise. Other changes occur, which include the altering of course content, adoption of a single textbook, establishment of time guidelines, and an increase in staff development. The most direct way to focus teachers' attention on tests is by establishing district objectives and publishing test results. Once teachers focus on test results, the most effective means of raising test scores is through a pretest.

Testing in Michigan

Researchers at the Michigan School Testing Service, Bureau of School Services, University of Michigan, have conducted a number of

studies to determine the nature and extent of testing in Michigan (Brezzezinski, 1976). Womer conducted the initial survey in 1958-59; he questioned members of the Michigan Association of Secondary School Principals regarding tests and test usage. The response rate was 50%. Twenty-two percent of the responding principals indicated that test results influenced inservice activities, communications with the public, and educational research. Seventy percent of the responding principals used the tests to identify exceptional children, for student placement, and for curricular evaluation (Womer, 1959).

Five years later, Womer (1981) conducted a second study, with the intention of drawing comparisons with the initial study. Ninety-three percent of the school districts responded to the inquiry. In 1963, the responsibility of developing a testing program shifted from the principal to counselors or other pupil-personnel specialists. This shift was reported by 50% of the respondents, whereas the percentage of counselors responsible in 1959 was 32%. The primary purposes, those of identification of exceptional children, student placement, and curricular evaluation, remained the same. There was a slight increase in the use of tests for research; this was offset by a small decrease in the use of tests for inservice and communications with the public.

Watson and Schmalegemeier (1970) conducted the third survey in 1968-69. The focus of this study differed from that of the previous two. The intention was to be "prescriptive" by revealing the interrelatedness of reported information. The response rate was

84%. "The most important use of test results is involved in the relationship between teacher and student" (p. 16). In addition, tests were for setting of educational goals and curricular evaluation. Organized testing programs were prevalent; 83% of the respondents had such a program.

The final study was conducted by the Department of Education in conjunction with the Michigan School Testing Service. Brezezinski designed two questionnaires; one was to gather extensive information from sampled districts, whereas the shorter version would be used to gather information from all districts. Response rates to the two questionnaires were 80% and 96%, respectively. What follows is a summary of the data from Brezezinski's study:

1. The primary person responsible for the testing program had shifted from the building principal to the counselor or pupil-personnel specialist.

2. Districts were using MEAP data in the following ways: Approximately 50% were reporting assessment results to the school board, and about 33% of the districts had the suggested curriculum study groups.

3. Approximately 50% of the teachers and principals did not believe they had the expertise to interpret results, nor could they apply criterion-referenced or norm-referenced results for student or curricular improvement.

In 1986, a joint study was conducted by the Office of Technical Assistance and Evaluation, Michigan State Board of Education, and

the Ingham Intermediate School District. All school districts were sent surveys, with a return rate of 79.8%. Ninety-eight percent of the respondents indicated that they were using MEAP results for the following purposes: to teach essential skills in a prescribed time frame, to report to the board of education, and to compare MEAP results to other test data.

In conclusion, these Michigan studies indicated that a majority of school districts had testing programs, the counselor or pupil-personnel specialist was responsible for programming and interpretation, and test results were used to identify exceptional children, program placement, and curriculum evaluation. Test results were being reported; however, staff personnel needed more training in test interpretation.

MEAP Studies

A number of studies have focused on the MEAP. The conception, development, and implementation of the MEAP have been well documented in both journal articles and doctoral dissertations. A brief summary of MEAP's history is provided in this section, followed by a discussion of several studies pertaining to MEAP.

Late in 1968, three staff members of the Michigan Department of Education's Bureau of Research began to discuss the lack of reliable information about the level, distribution, and progress of education. What resulted from these discussions was a proposal for a statewide educational assessment program. "Perhaps because of the rising concern about accountability in education, the idea received

support and endorsement of the state legislature and governor" (House, 1973, p. 48). It appears that the most important educational issue in the state was the inequitable distribution of school district achievement levels. Thus, the program would "provide members of the State Board of Education with information needed for allocation of the state's educational resources to equalize and improve the educational opportunities for all children" (House, 1973, p. 48).

The Michigan Public Acts of 1969 provided the initial funding for MEAP; subsequent funding was provided under Act 38 of the Public Acts of 1970. MEAP was part of a six-step educational management system. Those steps included the identification of common goals, development of performance objectives, assessment of educational needs, analysis of delivery systems, evaluation and testing of those systems, and recommendations for educational improvement. MEAP was designed to assess educational needs (Michigan Department of Education, 1972).

From MEAP's inception, the issue of whether to report the results publicly plagued the program. Legislators and state officials thought the public had a right to know, whereas superintendents, principals, and teachers thought that unjust comparisons would be made. However, the legislation did require that the Department prepare a ranking of all schools in the state. Further, they were to be prepared to release the point scores of all schools falling in the bottom quartile on two assessment measures-- composite achievement and relative socioeconomic level. Section III

monies of the School Aid funds were dispersed on the basis of this ranking in 1970-71 (House, 1973).

Jencka (1990) traced the evolution of MEAP in the following manner. Initially, the test was administered to all fourth and seventh graders in reading, math, and the mechanics of written language. This first test (1970) was norm referenced and remained so for the next four years. In 1973-74, two major changes were made: The test became a criterion-referenced test rather than a norm-referenced test, and the testing date was moved from January to September. The issuance of performance objectives sent a clear message to districts that these skills were what the state deemed important. The performance objectives were eventually replaced by "essential objectives."

Another revision in 1976 came in response to the criticism of "top-down" management. The reading and math tests were to be revised, based on the input of practitioners--those in the field. Also, all students in tenth grade were to be tested in the areas of reading, math, and the mechanics of written language. These tests were piloted in 1980.

In 1985, health testing began in the fourth, seventh, and tenth grades. Science testing in the fourth, seventh, and tenth grades was added in 1986; these tests were later mandated in 1988.

Revisions continued, and in 1989 a new reading test was mandated; science tests were administered in the fifth, eighth, and eleventh grades rather than the fourth, seventh, and tenth; and writing tests were piloted in the third, sixth, and ninth grades.

Edward Roeber, Supervisor of the Michigan Education Assessment Program, indicated that both math and science objectives and tests are in the process of revision. Thus, MEAP has evolved from a norm-referenced to a criterion-referenced test. It has changed with input from practitioners and is administered at a number of different grades and levels.

Early MEAP Studies

Numerous studies focusing on the MEAP have been conducted. One of the first studies was conducted by Fisher and Roth in February 1972. The researchers questioned K-12 superintendents about their general reactions to the use of MEAP in their local districts. Seventy-three percent of the superintendents responded to the questionnaire, indicating that the assessment program was having a positive effect in some districts. They also were sharing the results with their communities, despite the fact that they were neither supportive nor nonsupportive of the program.

The Michigan Association of Professors of Educational Administration and House, Rivers, and Stufflebeam conducted studies that are better known. The study conducted by the Michigan Association of Professors of Educational Administration (1972) focused on the process used to introduce the assessment program in Michigan, the relationship of assessment to accountability, and the testing instrument used to collect the data. Their summary was critical of the MEAP.

In 1973-74, the National Education Association and the Michigan Education Association contracted House, Rivers, and Stufflebeam to "assess the educational soundness and utility for Michigan of the Michigan Accountability System, with particular emphasis on the assessment component" (House et al., 1974, p. 1). These researchers viewed documents and gathered written and verbal testimony before rendering criticism of the accountability system. They reported that:

The most serious flaws are in the assessment component; it is too narrow in scope to serve as a state assessment, and it has been implemented on an every-pupil basis without technical or utility justifications. Overall, there is no evidence that state and local decision-making is being served by the accountability model. (p. 24)

These studies have provided information to practitioners and the Department of Education about the perceived problems associated with this assessment. The goal of these reports was to modify the program and in doing so "contribute to the improvement of education for all children in Michigan" (Aquino, 1975).

Dissertations on MEAP

Numerous dissertations are available, in which various aspects of the MEAP were studied. McCormick (1978) reviewed the stormy history of MEAP. Music (1982) investigated the test-textbook link in reading texts and found that although the Houghton-Mifflin series objectives were the most congruent with MEAP objectives, groups using this series did not perform better than the other two groups in her study. Bleacher (1975) concluded in his study of K-8 teachers that the assessment program was inconsistent with school

purposes, incompatible with their personal interests, and impossible to implement. Fidler (1988) indicated that the MEAP science scores were correlated with socioeconomic levels. Farfsing (1988) found that schools that scored higher on MEAP perceived themselves to be doing better than those that scored low.

Shinsky (1983) studied special education students and concluded that students from various disability groups scored differently on the MEAP reading and math tests. Simon (1982) found that teachers were good predictors of students' performance on MEAP; where teachers erred, it was in an underestimation of students' performance.

Several dissertations are more germane to the utility and use of MEAP results. Dissertations that are reviewed more extensively in the following pages are those by Aquino, Steele, Jencka, Thelander, and Bushaw.

The Aquino study. In 1975, Aquino questioned fourth- and seventh-grade teachers about their attitudes toward and use of the MEAP. A questionnaire was sent to randomly sampled teachers, with a response rate of 44% for fourth-grade teachers and 50% for seventh-grade teachers. Aquino found that assessment test results were received by 80% of the respondents.

Fourth-grade teachers believed they were "well prepared" or "very well prepared" to make use of test results; seventh-grade teachers felt less prepared. Approximately 75% of the fourth-grade teachers reported that the principal or assistant principal assisted

in the interpretation of test results; less than 50% of the seventh-grade teachers reported administrative assistance.

The primary use of the assessment instrument was to identify the instructional needs of groups of students; it was less helpful in identifying individual instructional needs and only somewhat useful in planning classroom instruction. Fourth-grade teachers tended to provide remedial help to students on nonmastered skills; seventh-grade teachers provided remediation less often.

Aquino found that only a small number of teachers were asked to change their instructional practices based on MEAP results. A small number of teachers reported being criticized because of MEAP scores. "The greatest impact of the assessment test results has been to confirm previous judgements teachers have had about individual student skills" (Aquino, 1975, p. 84).

The Steele and Jencka studies. Steele (1976) surveyed elementary and secondary principals in an effort to describe administrators' practices and attitudes toward MEAP. Seventy-five percent of the surveys were returned, with administrators indicating that they were using the assessment results in determining the general achievement levels of students and in determining the strengths and weaknesses of the curriculum. Extensive use of MEAP results was associated with building-level plans and the sharing of test results shortly after their arrival. The extensive use of these results was also associated with urban settings and districts with high minority and low-income populations.

Jencka (1990) replicated the Steele study to ascertain whether changes had occurred in the administrators' practices and attitudes toward MEAP. The response rate was 62%. Ninety-eight percent of the respondents used MEAP results to determine which essential skills were taught, to report to the board of education, or to use MEAP in conjunction with other test data. Administrators reported it was "quite useful" in diagnosing students' learning needs, planning instructional improvement, and communicating achievement levels of students and parents. As in Steele's study, extensive use of MEAP results was associated with schools that had building-level plans to target needs identified by MEAP.

In both the Steele and Jencka studies, the population sampled was elementary, junior high, and high school administrators. Jencka's analysis of data differed from Steele's in that, in addition to the administrative levels, he analyzed his data using the six strata identified by the Michigan Department of Education. Stratum six contains all districts in the northern Lower Peninsula and all districts in the Upper Peninsula. Jencka did not find significant differences across strata.

The Thelander study. Thelander (1979) compared the perceptions of local school board members, administrators, and teachers regarding the influence, use, and problems associated with MEAP. The sample was limited to Macomb County, Michigan. The response rate was approximately 93%. Results indicated that, although MEAP has influenced curriculum decision making, these board members, administrators, and teachers did not perceive the Department of

Education as being directly involved in local school district decision making. Respondents did not believe that state funds should be allocated on the basis of MEAP test results. They all indicated that teachers underused MEAP because of the lack of technical information and understanding; they believed this was due to few or poor inservice activities. All believed that MEAP could be expanded despite the perception that it did not meet district or individual student needs.

Thelander summarized his findings in the following manner: MEAP is useful in determining district goals, developing curriculum, and determining K-6 weaknesses. Instructional methods have been altered to improve MEAP results. However, "the usefulness of Michigan's present, statewide assessment must be questioned if it is not meeting the need of local districts" (Thelander, 1979, p. 116).

The Bushaw study. Bushaw (1988) randomly sampled all elementary school teachers and corresponding administrators based on demonstrated outcomes or lack of outcomes, as measured by the 1984-1986 MEAP results. Response rates varied from 41% of the teachers and 73% of the elementary principals to 80% of the superintendents. Perceptions of these three groups varied with respect to the existence of a policy statement; superintendents more frequently said that policy statements existed than did teachers. However, all three groups indicated that more actions occurred when policy statements were written; this action included the recognition of students who performed well, the establishment of a district

committee, the establishment of a goal for minimum attainment, and related inservice activities. Building-level activities included notification to the parents, the teaching of test-taking skills, and the establishment of attainment goals and a building-level committee.

What Bushaw found, however, was that a policy statement and actions based on that policy did not lead to changes in MEAP scores. "Teachers reported more policy statements and more district and building level actions in schools where MEAP scores had remained unchanged during the last three years than schools in which MEAP scores had improved significantly" (Bushaw, 1988, p. 114). In addition, MEAP was stressed to teachers, as indicated by 90% of the respondents. Seventy-five percent of the teachers changed instructional strategies as a result of reviewing MEAP. Said Bushaw, "It appears that the implementation of a policy to utilize MEAP results to improve student achievement as measured by the MEAP demonstrates symbolic rather than material consequences" (p. 123).

Summary

In conclusion, despite MEAP's stormy beginning, it has survived and continues in growth with respect to content and number of grades tested. The studies reviewed suggested that teachers and administrators were using MEAP test results. They reported using MEAP for the development of district goals, developing curriculum, determining strengths and weaknesses of their present curriculum, diagnosing student needs, and reporting to the public. Teachers

reported that they had altered their instructional methods to improve the students' test scores. However, teachers remained opposed to using these results for comparisons or for district evaluations.

Chapter Summary

In the 1960s and 1970s, federal and state governments became more involved in policy making that affected the local schools' educational programs. This involvement can be traced to the following five developments: (a) court involvement in equity issues, (b) the passage of the Elementary and Secondary Education Act, (c) the release of the Coleman Report, (d) the formation of the National Assessment of Educational Progress, and (e) the back-to-basics movement. These developments led to the loss of local control over educational programs and increased the reliance on standardized tests as measures of effectiveness.

The 1980s brought a continued movement toward accountability. Michigan's assessment program continued to expand with demonstrated use of its results. Originally intended to supply data to allocate resources, it now was intended to encourage local districts to review their curricula to ensure that all students mastered "essential objectives" in reading, mathematics, and science.

Several questions remain unanswered with respect to small, rural schools. Is MEAP being used to the extent reported by others? Is it being used for curricular development and review? Are students' needs being identified through the use of MEAP? Are MEAP

test results communicated to the public? The methods described in the following chapter were designed to answer these questions.

CHAPTER III

METHODS AND PROCEDURES

The Population and Sample

The population of this study consisted of small, rural school districts with total enrollments of 500 or fewer students. Further, these districts were located in the Upper Peninsula of Michigan. The list of small, rural school districts was generated from the Michigan Education Directory and Buyer's Guide (1990). At the time of this study there were 21 such school districts.

Ten districts were selected at random to participate in the study. The student enrollments of these ten districts ranged from 100 to 500. The districts were geographically distributed to include at least one district in each of these areas: the Copper Country, the western Upper Peninsula, the central Upper Peninsula, and the eastern Upper Peninsula. The districts were located in communities of economic decline, varying in degrees of economic hardship. The sample included districts experiencing extreme financial difficulties as well as those that were less concerned about their financial condition.

The selected districts had either a single facility or an elementary school and a high school combination. When driving through these communities, one could predict the location of the

school building--either on the main street or a block or two from the main street. Community involvement and pride in the schools was evident through the posting of events held at the school and signs commemorating athletic honors.

Of the districts surveyed, four districts were selected for extensive interviews. This selection was based on the following factors: degree of isolation, student enrollment, type of community, and academic history. These districts were selected since the researcher believed that they varied somewhat on these dimensions. This methodology was used to add depth to the survey of MEAP usage in small, rural school districts. Districts were assigned fictitious names to preserve anonymity. A brief summary of the districts in which interviews were conducted follows.

Aholahti had a student population of approximately 500. It was the largest school district in the survey sample. This district was adjacent to a large school district with a student enrollment of approximately 1,000. It had a long tradition of supporting education, with community members boasting of the number of graduates who were accomplished physicians, engineers, teachers, and attorneys. However, this district was located in a community undergoing severe economic decline.

Koskitown had a student population of approximately 350. In this study, it was of moderate size. It was surrounded by districts with student populations of approximately 1,000. Area educators recognized this district as one that supported continued formal education and inservice activities beyond what surrounding districts

were able to support. The community in which this school was located tended to be declining economically but maintained the hope of future mining activity.

Makivil was determined to be the most isolated as it was approximately 60 miles from the nearest adjacent school site. It also had one of the smallest enrollments in this study--approximately 100 students. Adjacent school districts had student enrollments of approximately 1,500. Because of its size and isolation, this district had experimented with various delivery systems. The district had been stable in terms of student enrollments but was starting to decline economically.

Vilmi enrolled approximately 125 students. Neighboring districts included one of similar size and one of approximately 700 students. This school was included in the original effective schools research because of its long-standing record of academic achievement despite its low socioeconomic status. Vilmi's student population was fairly stable; like other communities in the study, it was described as declining.

In summary, the population consisted of all K-12 public school districts in the Upper Peninsula of Michigan enrolling 500 or fewer students. Ten districts were randomly selected from the population to receive the MEAP survey. Of these ten districts, four were selected for more extensive study based on the factors of isolation, student enrollment size, type of community, and academic history.

The Survey Instrument

The survey instrument was designed to examine the integration of MEAP results into the existing structures of small, rural schools in the Upper Peninsula of Michigan. The instrument was designed to examine the perceptual responses of superintendents, principals, and third-, fourth-, sixth-, seventh-, ninth-, and tenth-grade teachers. These grade levels were selected since teachers at these levels may be teaching MEAP objectives or possibly providing remedial work based on MEAP test results. The unit of analysis was the collective responses of district personnel.

The general area of inquiry, that of the effect of Michigan's mandated assessment test, focused on the utility of the test results and how these results were used. The staff perceptions in each district were collectively compared to the staff perceptions in other districts regarding utility and usage.

Items on the survey instrument were developed after a careful review of literature, including writings on both small, rural schools and state assessment tests. Several items were developed from Michigan Department of Education publications; a majority of these items came from Pencils Down, A Guide for Using and Reporting Test Results (Michigan Department of Education, 1985). Some items were generated as a result of the concerns and uses of MEAP in this researcher's school district. In addition, one question was selected and replicated from Jencka's (1990) study of MEAP.

The questionnaire was modified several times after receiving input from educators participating in the field testing, officers of

Michigan's Rural Education Association, the supervisor of MEAP, and professors at Michigan State University. Fifteen staff members representing seven school districts participated in the field testing, with a 100% response rate. Input was solicited from Cleland Methner and Henry VerBerkmoes, president and past president, respectively, of the Rural Education Association in Michigan. Edward Roeber, supervisor of MEAP, Michigan Department of Education, also participated in the development of questions. Feedback at the university level was received from the researcher's doctoral guidance committee. Based on the recommendations from these four sources, as well as Dillman's (1978) Mail and Telephone Surveys: The Total Design Method, the survey was developed into its final form.

Collection of the Data

Before sending the surveys, this study was announced to the superintendents of Upper Peninsula schools at a conference held in Marquette, Michigan, on April 11, 1990. At this conference, the superintendent of the researcher's district requested help and cooperation in the completion of the surveys. Also, before sending the surveys, the Rural Education Association of Michigan was contacted, and this researcher met with the governing board to discuss the study and solicit cooperation. In addition, the U.P. Outreach Director of the Michigan Department of Education contacted the selected schools to encourage cooperation in this study.

Superintendents in the districts to be surveyed were asked whether their districts would participate in the study.

Individuals, in most cases the superintendent's secretary, were contacted in April and asked if they would distribute and collect the surveys. Packets were sent to these identified individuals. These packets contained a cover letter indicating the intention of the survey, a list of individuals to receive the survey, a set of directions for distribution and collection, and the individual survey packets. According to instructions, individuals would receive a cover letter, a survey, and an envelope for the completed survey. Voluntary participation was emphasized in the cover letter. Telephone calls were made to the districts a week after the anticipated receipt of the packets.

The researcher went to each of the ten districts to collect the completed surveys. Brief conversations occurred in all surveyed districts regarding discussions generated about this inquiry or district comments about MEAP.

The survey response rate is shown in Table 1. The response rate for nine of the ten districts was excellent. However, the response rate from one district was lower than desired. Several contacts were made with this district, and additional surveys were sent to the district's elementary school. The requirement of anonymity, coupled with the conclusion of the school year, prohibited a district visit and individual requests of nonrespondents.

Table 1.--Number of surveys sent to and received from participating districts.

District	Surveys Sent	Surveys Returned	Percent
1	13	13	100
2	9	8	89
3	14	14	100
4	12	12	100
5	7	7	100
6	6	6	100
7	11	7	64
8	11	11	100
9	8	8	100
10	10	10	100

After the surveys were collected, appointments were made with the superintendents and principals in the four districts in which interviews would be conducted. Interviews were held with both administrators and teachers. Teacher interviews were not limited to those surveyed. In all cases, if a counselor was employed by the district, this individual was also interviewed. The interviews were based on two basic questions: "How do you know you are teaching what you ought to teach?" and "Where do standardized tests fit in your classroom program?" Interviews were tape recorded, with permission granted by each interviewee before the taping. Interviews were conducted during the first three weeks of May. In the three larger districts, at least ten individuals were interviewed. In the fourth district, all available administrators and teachers were interviewed.

Data Analysis

Data from the returned surveys were entered into computer files by districts. Data analysis was accomplished using Statistical Analysis Systems (SAS, 1985).

A majority of the research questions were analyzed using analysis of variance (ANOVA). Questions 4b, 4c, 4d, 5, 6, 7, and 8 were analyzed using one-way ANOVA, comparing district means with other district means in the sample for each question. A post-hoc test, the Scheffe, was done when differences occurred. The criterion for statistical significance was set at the .05 alpha level.

Questions 3 and 4a were reported as the proportion of the districts selecting "yes" and the proportion of the districts selecting "no." Question 1 was reported as the proportion of districts that "strongly agree," "agree," "neutral," "disagree," or "strongly disagree" with the use of standardized tests as measures of district effectiveness. Question 2 was reported as the proportion of the districts using MEAP results in conjunction with other standardized tests, the proportion of districts that used only MEAP, and the proportion of districts that did not use MEAP.

Response rates varied from question to question, based on the coded-in response of those surveyed. Questions with fewer responses included 4, 6c, and 8c. These questions pertained to the availability of curriculum guides, the discussion of MEAP results at school board meetings, and parents' knowledge of their students' MEAP scores.

Finally, interviews were conducted in four selected districts to develop further the areas of inquiry. Two basic questions were asked of administrators and teachers: "How do you know you are teaching what you ought to teach?" and "Where do standardized tests fit in your classroom program?" The responses to these interviews are reported as case studies and are also woven into the fabric of Chapter IV.

CHAPTER IV

PRESENTATION OF DATA

Introduction

The researcher's purpose in this study was to examine the effect that MEAP has had on the curriculum in small, rural school districts. The effect was measured by investigating the presence of MEAP in the testing program, in curriculum decisions, in classroom teachers' content decisions, and in the discussion of test results. In the first section, percentages are presented to indicate the educators' beliefs about the usefulness of standardized test information, the use of MEAP in conjunction with other tests, and the presence of curriculum guides. The second section contains the mean ratings and discussion of rural educators' responses to Questions 4b through 8d. This section includes a discussion of the use of MEAP test results, the communication of MEAP test results, and the changes that have occurred based on the districts' MEAP test results. In the third section, four vignettes are presented to describe briefly to the reader the types of school districts surveyed. In all sections, interview material is incorporated to substantiate or extend the data collected in the survey.

From the data collected, responses were sought to the following research questions:

1. Are there significant differences among districts with respect to the use of MEAP test results?
2. What practices, if any, are intentionally used by small, rural schools in response to the MEAP?

Survey Returns

On April 22, 1990, 101 surveys were sent to superintendents, principals, and third-, fourth-, sixth-, seventh-, ninth-, and tenth-grade teachers in the districts selected for inclusion in the study. A majority of the surveys were collected by May 7. The outstanding surveys were received in the first few weeks of June. One district, because of its low return rate, received second copies of the survey in late May in an effort to increase the percentage of educators responding from that district. The total number of surveys returned as of July 1, 1990, was 94. Thus, the response rate to this survey was 93%. On several questions, the response rate was somewhat lower due to omissions.

Interviews

Interviews were conducted in Aholathi, Koskitown, Makivil, and Vilmi during May. The two interview questions asked were as follows: "How do you know you're teaching what you ought to teach?" and "Where do standardized tests fit into your classroom program?" Interviewees in each district included an administrator, a counselor, a compensatory education teacher, and teachers of all grade levels. Teacher interviews were not restricted to those completing the survey. In Makivil, all teachers were interviewed.

In the remaining districts, interviews were sought from a teacher at each grade level.

Results of Data Analysis for Test Usage

In this section, the collective percentage of rural educators across their districts is reported per item. The intention of these items was to determine the perceived use of standardized tests in small, rural schools as well as the perceived use of MEAP test results in relationship to standardized tests. In the following paragraphs, the item is presented along with the interview material related to that item.

Questionnaire Item 1

Standardized tests (SAT, ACT, Stanford, Metropolitan, MEAP, etc.) provide important information in assessing your district's effectiveness.

The responses of district administrators and selected teachers to this item are shown in Table 2. Educators in small, rural districts did not represent the entire spectrum of beliefs about the usefulness of standardized tests for assessing effectiveness. No one "strongly disagreed" with the use of tests for this purpose. Respondents from some districts confined their perceptions to two categories, whereas others used a broader range, spreading responses across four categories.

Table 2.--Responses to Item 1: Standardized tests provide important information in assessing your district's effectiveness, by individual districts.

District	Strongly Agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly Disagree (%)
1	30.8	38.5	30.8	0.0	0.0
2	25.0	50.0	12.5	12.5	0.0
3	7.1	42.9	35.7	14.3	0.0
4	0.0	63.6	18.2	18.2	0.0
5	14.3	28.6	42.9	14.3	0.0
6	33.3	66.7	0.0	0.0	0.0
7	16.7	66.7	16.7	0.0	0.0
8	9.1	54.5	18.2	18.2	0.0
9	12.5	87.5	0.0	0.0	0.0
10	11.1	66.7	11.1	11.1	0.0

In this section, tests are referred to using the following acronyms: Preliminary American College Test (PACT), Preliminary Scholastic Aptitude Test (PSAT), American College Test (ACT), Scholastic Aptitude Test (SAT), Armed Services Vocational Aptitude Battery (ASVAB), and Science Research Associates (SRA).

Respondents from Aholathi, reported as District 1 in Table 2, agreed that tests provided information used to measure the effectiveness of their district. Interviewees, however, painted a slightly different picture. Interviewees from this district consistently reported that its success or effectiveness was demonstrated by the percentage (90%) of students going to college and the high percentage of successful physicians, attorneys, engineers, and teachers. Other indicators included an almost nonexistent dropout rate--one every three or four years, a higher

average on the ACT than the national average (22 versus 18), an attendance rate of 95%, and the lack of serious discipline problems.

Less tangible measures of success in Aholathi included the community's satisfaction with the educational program, the active involvement of parents in school activities, and the number of happy students. Teachers in this district found that student feedback best measured their effectiveness and provided direction in instructional content.

Aholathi respondents' indecisiveness on the survey was indicative of the staff members' struggle to add the Stanford Achievement Test to their testing program; some members were eager to adopt the test, whereas others wanted nothing to do with test results. With the exception of the reported ACT statistic, probing questions were required to evaluate the importance of tests in this district. The typical response, with the exception of the superintendent, was that tests were not very important. However, some respondents indicated a desire for more tests and that tests become more important.

Interviewees from Koskitown, reported as District 4, verbally indicated that tests were fairly important in the measurement of success. They looked seriously at the ACT, the California Achievement Test, and the MEAP. Comparisons were made with previous years on these tests. Staff members met specifically to discuss test results and develop plans to compensate for perceived weaknesses. Yet, in this district, some members preferred to ignore test

results, generating what was important for their students from experience, readings, and other experts. Overall, staff members felt pressured to "show well" on tests.

However, staff members in Koskitown also measured effectiveness by the atmosphere of the school, the support of the community, the number of students qualifying for scholarships, and the number of students attending college. Other factors that indicated success in Koskitown were the low dropout rate, the high attendance rate, and high ACT scores. Koskitown's freshmen had a higher grade point average than other freshmen at Northern Michigan University. Data presented in Table 2 tended to be confirmed by the interviews in Koskitown (District 4). Tests provided important information about their effectiveness. However, because of the degree to which tests were used, some resisted their use; this is congruent with the findings from the Darling-Hammond and Wise (1985) study.

Respondents from Makivil, represented by District 5 in Table 2, tended to agree that tests provided important information about their district's effectiveness. Responses in interviews also ranged from the use of tests in math and physics competitions to "I don't pay much attention to them." However, overall, testing played a very small part in this district; observation played a larger role. Said one staff member, "I have these students for three hours each day, and I know what they know and don't know." However, as reported by the superintendent, success was demonstrated by high test scores.

Respondents from Vilmi, represented by District 9 in Table 2, reported that standardized tests provided useful information in the assessment of their district's effectiveness. Interviewees indicated that tests were used to determine whether the district was doing the job, but "certainly tests do not have a major impact on altering the course of change." In general, staff members said, "We do well," but, in the same breath, they indicated that "students don't take the tests seriously enough." Staff members distrusted test results and indicated that these results underrepresented the skill levels of their students. Vilmi personnel relied on feedback from their alumnae to measure their success. They also measured their effectiveness by the number of students going to college, their high attendance rate, their low dropout rate, and their daily interactions with students.

The percentages for each item response in Question 1 are collapsed across districts and presented in Table 3.

Table 3.--Percentages for each item response in Question 1--Standardized tests provide important information in assessing your district's effectiveness--across all surveyed districts.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
All districts	15.1	54.8	20.4	9.7	0.0

Respondents from small, rural districts tended to agree that tests provided important information in assessing their districts' effectiveness. Approximately 70% of those surveyed agreed or strongly agreed with the use of tests in this manner. Of those responding to the survey, less than 10% disagreed with the use of tests as a measure of effectiveness. In conclusion, respondents from small, rural districts in the Upper Peninsula of Michigan tended to be favorably disposed to the use of test results to measure success in their school districts.

Questionnaire Item 2

The following most accurately describes our district:

MEAP results are used in conjunction with other standardized tests.

MEAP results are the only outcome measures used in our district.

MEAP results are not used.

The participants' responses to Item 2, the integration of MEAP results in the district's testing program, are shown in Table 4. Responses indicated that all combinations were presented in these Upper Peninsula school districts. However, in the majority of districts, MEAP results were used in conjunction with other standardized tests.

Table 4.--The use of MEAP results with other tests, MEAP alone, or not at all by the individual districts.

District	MEAP and Other Tests (%)	MEAP Alone (%)	MEAP Not Used (%)
1	69.2	23.1	7.7
2	75.0	0.0	25.0
3	100.0	0.0	0.0
4	90.9	9.1	0.0
5	71.4	14.3	14.3
6	100.0	0.0	0.0
7	100.0	0.0	0.0
8	70.0	10.0	20.0
9	100.0	0.0	0.0
10	90.0	10.0	0.0

Respondents from Aholathi, represented by District 1 in Table 4, generally indicated that MEAP was used in conjunction with other standardized tests. The tests administered in Aholathi were those for college or armed service entrance in addition to the MEAP. Approximately one-fourth of the teachers indicated the MEAP was used alone, which probably represented the elementary teachers' responses. Elementary teachers did not have another standardized test in place. Other individuals indicated that "we always do well, so I discontinued looking at them," "MEAP is used only for special education kids, I don't use it," and "I could use [standardized] tests, but that is not likely." Aholathi had a fair share of professionals believing that tests should measure the "general factor" and that students should not be taught the test's content.

In contrast with Aholanti, 90% of the teachers from Koskitown reported using MEAP in conjunction with other tests. Teachers in Koskitown "know that tests are important. . . . They do not want to have low scores." Said one staff member, "Students don't see them [tests] as important, teachers do. That's how we get evaluated. The reality of life is that no matter what is said to the contrary, what gets printed in the newspaper is how you are evaluated." Teachers were not opposed to teaching test objectives, matching the test question format, or teaching the test's vocabulary. Teachers indicated that they thought the administrators used test results to evaluate them. Koskitown administered the California Achievement Test in grades 1 to 8, the PACT in tenth grade, and the ACT and ASVAB in eleventh and twelfth grades. It might be speculated that the one individual who did not recognize the use of MEAP results might have been registering a protest vote.

In Makivil, represented by District 5 in Table 4, a majority of staff members responded that MEAP was used in conjunction with other tests. One staff member indicated that it was the only test used, whereas another staff member indicated that it was not used at all. Teachers in Makivil administered the Stanford Achievement Test in elementary school; the PSAT, PACT, ACT, and ASVAB were administered in the tenth, eleventh, and twelfth grades. MEAP was administered in the designated grades. One teacher administered math and chemistry tests that are developed nationally or statewide. Test results tended to be important if the recipient of the results valued the test, as in the situation of the math/science teacher.

This teacher said, "My math and chemistry tests provide useful information about the content of my courses, but tests and objectives, they should go away." Thus, in Makivil, tests tended to be used in conjunction with others, but this was not a unanimous opinion.

In Vilmi, represented by District 9 in Table 4, educators unanimously reported that "MEAP results are used in conjunction with other standardized tests." Vilmi had administered the Science Research Associates (SRA) test to K-12 students for a number of years. Other tests administered were the PACT, ASVAB, ACT, and SAT at the high school level. MEAP was given in the designated grades.

Responses to Item 2 were collapsed across districts to indicate the manner in which small, rural schools in the Upper Peninsula of Michigan used tests. These data are presented in Table 5. A majority of respondents, 86%, indicated that MEAP was used in conjunction with other tests. In only one of the schools surveyed, the elementary school in Aholathi, was a combination of tests not available to educators. Approximately 7% of the participants indicated that MEAP results were not used. However, as a whole, educators indicated through the survey instrument and interviews that MEAP results were used.

Table 5.--The use of MEAP with other tests, MEAP alone, or not at all across all districts.

	MEAP and Other Tests (%)	MEAP Alone (%)	MEAP Not Used (%)
All districts	86.0	7.5	6.5

Questionnaire Item 3

Our district selects or develops tests for items in the local curriculum but not measured by MEAP.

Responses to Item 3 are presented in Table 6 by individual districts.

Table 6.--Responses to Item 3: District selects or develops tests for items included in the local curriculum but not measured by MEAP.

District	Yes (%)	No (%)
1	15.4	84.6
2	50.0	50.0
3	46.2	53.8
4	72.7	27.3
5	42.9	57.1
6	83.3	16.7
7	50.0	50.0
8	66.7	33.3
9	25.0	75.0
10	60.0	40.0

Study participants tended to be polarized in responding to this question. Those from Aholanti, represented by District 1 in Table

6, tended to respond to the survey as they did in the interviews. The testing program was not well coordinated, and a need for more tests was indicated. Koskitown's staff responded as anticipated. A majority of educators indicated that tests were selected on the basis of goals and objectives. Koskitown's teachers felt pressured to teach the local curriculum and were perceived to be evaluated by the results of the selected tests. The staff in Makivil responded in a divergent fashion; however, instructors in Makivil were also divided on the issue of test use for measures of effectiveness and were not united in the reporting of their testing program. Vilmi's staff presented a united front in describing their testing program and were also favorably disposed to testing as part of a measure of effectiveness. Thus, Vilmi's response was anticipated and confirmed by interviews.

As in the previous items, the participant responses in this survey were combined to achieve a percentage rate across districts (see Table 7). Small, rural schools in the Upper Peninsula were almost evenly divided on the selection or development of tests for items not measured by MEAP. Individual school district profiles indicated that some districts selected additional tests to meet curriculum needs whereas other districts did not.

Table 7.--Percentages for each item response in Item 3--District selects or develops tests for items included in the local curriculum but not measured by MEAP--across all districts.

	Yes (%)	No (%)
All districts	47.3	52.7

Based on the data presented for the first three items, educators from small, rural districts tended to view tests favorably as a measure of effectiveness and tended to use MEAP results in conjunction with other tests. However, the districts collectively were split on the selection and development of tests for items included in the local curriculum but not measured by the MEAP.

Questionnaire Item 4

Curriculum guides are available in reading and math.

This question was presented in such a manner that respondents acknowledged either the presence or absence of such documents. Table 8 contains data for individual districts, whereas Table 9 includes data as a combination of responses across the ten districts participating in the study.

Table 8.--Responses to Item 4: Curriculum guides are available in reading and math.

District	Yes (%)	No (%)
1	53.8	46.2
2	62.5	37.5
3	64.3	35.7
4	63.6	36.4
5	0.0	100.0
6	33.3	66.7
7	50.0	50.0
8	54.5	45.5
9	25.0	75.0
10	80.0	20.0

In Aholahti, District 1 in Table 8, approximately 50% of the respondents indicated that curriculum guides were available, and approximately 50% indicated that the guides were not available in reading and math. Aholahti's high school was accredited by the North Central Association in the early 1980s but had subsequently lost this accreditation due to the dilapidated school plant. Guides were in place for the high school but hardly so in the elementary and junior high schools. Most teachers indicated no coordination across grades on curricular issues. One staff member angrily stated, "We have a curriculum committee here comprised of administration, board members, and parents--there is no teacher input." Another staff member indicated that "We have no curriculum; each teacher does their own thing. We don't even have staff meetings." These questionnaire responses were not surprising, given the lack of coordination in the elementary facility; this

percentage, 46.2%, roughly corresponded to the number of teachers in this facility. The percentage of staff who indicated that guides were available approximated the percentage of staff teaching in the previously accredited high school.

A majority of the staff members in Koskitown (District 4 in Table 9) indicated that curriculum guides were available in math and reading. Teachers, in interviews, talked about teaching the district's objectives and communicating these goals and objectives to students so that they would know what was expected of them. However, some teachers indicated that instruction was left to the teacher, with some guidance by the administration. At the time of this study, Koskitown did not have a valid elementary reading curriculum in place due to a stalemate by teachers and administration regarding the time and manner in which this new guide was to be generated. Elementary teachers in Koskitown had asked to receive a stipend for work to be done after the completion of the 1989-90 school year, and such a stipend was not funded. Thus, teachers had incorporated bits and pieces of a "literature-based" reading program without specified goals and objectives. This could account for a large portion of the participants responding to the lack of a curriculum.

Makivil respondents (District 5 in Table 8) indicated that curricula were not available in reading and math. Considering that this district employed ten teachers, four teaching in the elementary school and the remainder in the high school, guides could be

generated by a single teacher. In this district, the lack of guides could be a function of time, not a function of time and communication as in other districts.

Vilmi (District 9 in Table 8) had a North Central Association accredited high school; thus, guides were available. Approximately 25% of the staff in Vilmi responded that guides were available; this was the same proportion of high school teachers to instructional staff. Elementary teachers did not refer to goals, objectives, or guides when interviewed; the researcher assumed that guides were not used.

Based on the data presented, some of the surveyed districts had curriculum guides in place; others did not. The guides tended to be in place at the high school level because of accreditation standards. As shown in Table 9, districts were quite evenly divided on the presence of guides in reading and math.

Table 9.--Percentages for each item response in Item 4--Curriculum guides are available in reading and math--across all districts.

	Yes (%)	No (%)
All districts	52.1	47.9

Questionnaire Items 4b, 4c, and 4d

MEAP objectives are incorporated into our district curriculum guides.

Textbooks are chosen on the basis of our curriculum guides.

Curriculum guides influence classroom instruction.

Only those respondents who indicated that their district had curriculum guides answered Items 4b, 4c, and 4d. Thus, the number of responses for these three questions was 48 each. Data for these three items are shown in Appendix Table D.1.

Item 4b. MEAP objectives are incorporated into our district curriculum. The mean score for this item was 3.92, indicating that, on the average, the total group perceived that MEAP objectives were "somewhat" incorporated into curriculum guides. Thus, if curriculum guides were present in the district, MEAP objectives would be found in these guides. However, these objectives were not "extensively" found in the guides because the MEAP objectives tended to be viewed as "minimal." These guides would include objectives designed to extend the skeletal framework provided by MEAP objectives.

Item 4c. Textbooks are chosen on the basis of our curriculum guides. This item had a mean rating of 4.10, indicating that, on the average, the total group perceived textbooks to be chosen "somewhat" on the basis of the curriculum guides. Statistically significant variability was not found among districts, with an F-value of .058, but the interview component of this study tended to disclose variability.

Aholahti respondents, who remained fairly divided on the presence of guides, also reported that there was very little coordination of curriculum. Elementary and junior high school teachers indicated that they individually made a selection and that selections were "not in a series. In our method, there is a lot of duplication."

In Koskitown, quite a different picture was presented. The superintendent in Koskitown indicated that they were consciously trying to follow the process of developing the curriculum first, with textbook adoption to follow. The district had experienced disappointments when the process was reversed, as in the selection of a math book about ten years before. The selection of the sixth-grade social studies book was based on a curriculum that integrated MEAP objectives with local objectives.

Item 4d. Curriculum guides influence classroom instruction.

The mean rating for this item was 4.36, indicating that, on the average, educators in small, rural districts allowed guides to influence class instruction "some." Koskitown teachers' instruction was based on "goals and objectives . . . the state model and what our curriculum indicates," as stated by the superintendent. Koskitown's principals agreed, saying, "Teachers make a conscious effort to teach objectives." The Koskitown teachers' responses paralleled those of administrators in the district: "We have a written curriculum and we adjust as we go . . . do modifications based on the kids."

Teachers' instruction in Aholathi, Makivil, and Vilmi tended to be less influenced by guides. One Makivil teacher said, "Textbooks guide my teaching. I would be helter skelter without them." A fellow teacher in Makivil followed with, "I use the textbook and follow that curriculum." Teachers in Aholathi and Vilmi tended to use the textbook as the guiding document, with final decisions based on the availability of materials and personal experience.

In summary, districts did not differ in a statistically significant manner with regard to the incorporation of MEAP objectives in curriculum guides, selection of textbooks based on curriculum guides, or the guides' influence on classroom instruction. However, based on interview responses, Koskitown appeared to have a closer alignment of guides to textbooks and of textbooks to instruction.

Use of the 1989 MEAP Results

Educators in the sampled small, rural schools were asked to respond to ten statements regarding the extent to which they used the 1989 MEAP test results. The extent to which rural educators used these results for each of the selected purposes is discussed in the following paragraphs. In addition, if districts varied significantly, those contrasts identified are discussed. However, other significant contrasts may be present but have not been identified. Mean scores for these test items are reported in Appendix Table D.2.

- a. To determine the general achievement level of tested students.

The mean rating for this purpose was 3.96, indicating that, on the average, rural educators used MEAP results to "some" extent to determine the general achievement level of tested students in their schools. This appeared to be one of the more important purposes, ranking somewhat lower than determining the strengths and weaknesses of the math and reading curricula.

- b. To inform the school community of the general achievement level of tested students.

Rural educators gave this purpose a mean rating of 3.87. These educators indicated that, on the average, they made "some" use of the 1989 MEAP test results to inform their respective communities of their students' general achievement level. Of the four districts in which interviews were conducted, none had a newspaper published in its community. Aholathi and Vilmi could report their results to a bi-county newspaper; however, residents in Vilmi typically did not subscribe to the newspaper. Neither Koskitown nor Makivil received newspaper coverage. Koskitown shared its results as an insert in a "Buyer's Guide," and Makivil shared its results in a school letter sent to all box holders. Comparisons on achievement levels of students from district to district were less direct due to the lack of newspaper coverage.

- c. To determine the strengths and weaknesses in the area of math.

The mean rating for this purpose was 4.14, which indicates that, on the average, rural educators used the 1989 MEAP results "some" to determine the strengths and weaknesses in the area of math. The MEAP test results were used by staff members in this survey to indicate the adequacy of their curricular content. Interviews indicated that if MEAP results were reviewed, districts most frequently reported weaknesses in the fraction and decimal area. Teachers in Aholathi, Koskitown, and Vilmi emphasized these math concepts, insuring that the textbook material was covered instead of deleting it, as was typical at the end of the school year.

- d. To determine the strengths and weaknesses in the area of reading.

This purpose received a mean rating of 4.16, the highest mean recorded in this section on test-result use. This rating means that, on the average, educators in this study used the most recent MEAP results "some" to determine the strengths and weaknesses of their reading curriculum.

Reading had been in the forefront of curricular issues in Michigan due to the introduction of a new definition of reading. The 1989-90 MEAP reading test was the first assessment of reading based on this definition. District educators were eager to review results as the Michigan Department of Education predicted that two out of three students would fail this assessment test. However,

interviews did not reveal any particular concern about the reading curriculum, even with this new definition.

One district, Vilmi, had previously had concerns in the reading area--particularly about the objectives regarding prefixes. The instructor in Vilmi said, "We just made sure that instruction on prefixes was included the next year."

e. To determine instructional priorities.

Educators in small, rural schools gave this purpose a mean rating of 3.87, indicating that, on the average, they used MEAP results "some" to determine instructional priorities. However, rural educators differed significantly on this purpose, as indicated in Table 10.

Table 10.--Results of analysis of variance for Item 5e: To determine instructional priorities.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Instructional priorities	9	63.283	7.031	2.09	.039*
Error	83	279.168	3.363		
Corrected total	92	342.452			

*Significant at the .05 alpha level.

Interviews did not elicit any particular comments on this subject. What educators tended to mention most was a concern about

inclusion of subject matter, but the prioritizing of subject matter was not seen as an issue worthy of comment. As indicated in Table 10, respondents from the various districts differed with regard to the use of MEAP results to determine instructional priorities.

f. To determine placement of students in "remedial" programs.

The mean rating for this purpose was 2.82. This indicates that, on the average, administrators and teachers used MEAP results "very little" to determine the placement of students in "remedial" programs. However, the collective responses of educators by their respective districts differed significantly (see Table 11). The following contrast was found using the Scheffe test: Districts 2, 7, 8, 10, and Vilmi differed significantly from Aholahti, Koskitown, and District 6 on the use of MEAP test results to determine placement of students in "remedial" programs. Respondents from Aholahti, Koskitown, and District 6 indicated that they used these results "some," whereas the other identified districts used the results "very little." The mean ratings were 1.8032 versus 3.6587 for this use.

Table 11.--Results of analysis of variance for Item 5f: To determine placement of students in "remedial" programs.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Student placement in remedial programs	9	94.671	10.519	2.47	.015*
Error	82	349.187	4.258		
Corrected total	91	443.859			

*Significant at the .05 alpha level.

Aholahti's compensatory teacher indicated that MEAP test results were included in the needs assessment for students considered for the program. MEAP test results were not used in instructional planning. Vilmi and Makivil educators did not use MEAP results in a methodical way for their "remedial" or compensatory education programs.

Koskitown, however, in the past had established a "remedial" program, typically in math, for those students who did not successfully pass the test. Seventh and tenth graders were required to take a special math class, which stressed the mastery of math objectives. Students could test out of the class at any time. This program was unique to those schools surveyed and interviewed.

g. To determine the need for new programs.

For this purpose, rural educators gave a mean rating of 2.76. On the average, educators used MEAP results "very little" to

determine the need for new programs. Comments by rural educators in interviews indicated that rural schools in the Upper Peninsula of Michigan offered only the basics. They were struggling to maintain the basics; adding new programs was out of the question.

h. To analyze teacher performance.

The mean rating for this purpose was 1.521, indicating that, on the average, staff members in rural schools perceived that MEAP results were used "very little" to analyze teacher performance. This purpose received the lowest mean rating of the ten purposes presented. Considerable discussion was generated on this purpose in teacher interviews, particularly in Koskitown.

i. To identify staff-development needs.

Rural educators gave this purpose a mean rating of 2.591. This means that, on the average, MEAP results were used "very little" to identify staff-development needs. However, respondents from the various districts differed significantly on this purpose (see Table 12). District 10 differed significantly from Aholanti, Makivil, and Vilmi on this dimension. This was the only contrast found, with mean ratings of 4.8000 versus 1.6230. District 10 used MEAP "quite a bit" to identify staff-development needs, whereas other districts used MEAP "very little" for this purpose. Interviews were devoid of comments on this subject.

Table 12.--Results of analysis of variance for Item 5i: To identify staff-development needs.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Identify staff-development needs	9	86.975	9.664	2.56	.012*
Error	83	313.498	3.777		
Corrected total	92	400.473			

*Significant at the .05 alpha level.

j. To predict students' future academic success.

For the purpose of predicting students' future academic success, rural educators gave a mean rating of 2.398, indicating that, on the average, MEAP results were used "very little." Interviews did not add any insight into this subject.

Comparison With Jencka Study

Item 5, the use of MEAP results for specified purposes, was a duplication of an item used in a statewide survey of principals (Jencka, 1990). The mean rating of this study is compared with that of the Jencka study in Table 13.

Table 13.--Mean ratings of the use of MEAP results for specified purposes by rural educators and by principals in Michigan.

Purpose	Mean Rating	
	Rural Educators	Principals
General achievement level	3.97	4.89
Inform community	3.87	4.72
Strengths and weaknesses in math	4.14	5.57
Strengths and weaknesses in reading	4.16	5.45
Instructional priorities	3.87	4.71
Placement in remedial classes	2.82	3.01
Need for new programs	2.76	3.15
Analyze teacher performance	1.52	1.87
Identify staff-development needs	2.59	3.38
Predict students' future academic success	2.40	2.71

Data gathered from this sample of educators in rural schools indicated that, overall, rural educators tended to use the MEAP results to a lesser degree than did the principals in the statewide sample in Jencka's study. However, this result aligns with the literature of Bushaw (1988), who indicated that administrators tend to place greater value on tests than do teachers. Interviews in Aholathi, Koskitown, Makivil, and Vilmi added credence to this observation. Administrators of these schools more frequently quoted test results or indicated that tests determined what content should be taught.

However, similarities were present in these two studies. The ranking of the reported means tended to be the same. Rural educators used MEAP results to detect strengths and weaknesses in

reading more so than in math. In Jencka's study, the order of these two purposes was reversed. This difference could be due to the inordinate amount of publicity surrounding the MEAP reading test in 1989-90, the first year in which this test was administered to all fourth-, seventh-, and tenth-grade students.

"Determining the general achievement level of tested students" was the third-ranked purpose in both studies. Parallelism of means continued until the sixth ranking, where rural educators rated "Placement of students in remedial classes" higher than did Jencka's principals, who ranked "Identify staff-development needs" in sixth position. Similarities of ranks continued through the remainder of the question.

In conclusion, the replication of this question indicated that although the means varied somewhat, similarities existed in overall rankings.

Summary of the Use of MEAP Results

On the average, rural educators tended to use MEAP results to determine curriculum strengths and weaknesses and to determine the general achievement level of tested students. They made "some" use of MEAP results to inform the school community of their students' achievement levels and to determine instructional priorities. The respondents in this study used MEAP results "very little" for placing students in remedial programs, determining the need for new programs, analyzing teacher performance, identifying staff-development needs, and predicting students' future academic success.

Questionnaire Item 6

The intention of this item was to assess the local school district's communication of MEAP test results. This assessment was confined to interpretation of MEAP results, discussions of MEAP results at district-level meetings, and discussions of strategies to incorporate MEAP objectives into the local curriculum. The extent to which small, rural school district educators communicated MEAP results is discussed in the following paragraphs. If districts varied significantly, a Scheffe test was used to identify significant contrasts. These contrasts are discussed; however, other contrasts might be present and significant but remain unidentified.

- a. The principal provides leadership in the interpretation of MEAP test results.

The mean response for this item was 2.974. Rural educators, on the average, indicated that the principal provided "some" leadership in the interpretation of MEAP test results. Respondents from the various small, rural districts varied significantly on this item (see Table 14). Using the Scheffe test, Makivil and Vilmi were found to vary significantly from the remaining districts. Respondents from Makivil and Vilmi (mean = .4643) indicated that the principal provided "very little" but closer to no leadership in the interpretation of test results, whereas those from the other districts indicated that the principal provided "some" leadership on this task (mean = 3.4484). This variability was supported by both administrator and teacher comments.

Table 14.--Results of analysis of variance for Item 6a: The principal provides leadership in the interpretation of MEAP test results.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Principal provides leadership in interpretation	9	213.993	23.777	5.98	.0001*
Error	84	333.837	3.974		
Corrected total	93	547.830			

*Significant at the .05 alpha level.

Comments by administrators are presented together to preserve anonymity of individual principals. In one district, the administrator said, "Don't ask me any questions about MEAP; see the counselor for those." In another district, administrators accepted the responsibility for the interpretation of MEAP results and conveyed these results at teacher meetings. The superintendents in the remaining two districts took the responsibility for communicating results. In these districts, communications were to the curriculum committee or to the school board. Interpretation to the teachers was minimal or nonexistent.

Teachers' comments substantiated the administrators' reports. Said one teacher, "I don't know if I should be doing more on some objectives, if I should be pulling back on others. . . . We do all right. Praise goes to the superintendent at the school board

meetings." Another teacher asked about standardized measures: "What does the number tell us? We don't do anything specifically with the test results."

Thus, at least one of the districts in which interviews were conducted had a principal who was actively involved in the interpretation of MEAP results. This district, along with other surveyed districts, provided the contrast to Makivil and Vilmi.

b. MEAP has been discussed at teacher meetings.

Small, rural district educators in this study gave this item a mean rating of 4.352, indicating that, on the average, staff members discussed MEAP results "some" at teacher meetings. Districts varied significantly on this type of communication (see Table 15). A significant contrast existed between Makivil and Vilmi (mean = 2.4732) and the remaining districts, excluding Aholahti (mean = 4.8676).

Respondents from Makivil and Vilmi indicated, through the survey, that MEAP had been discussed "very little" at teacher meetings. Educators from the remaining districts indicated that MEAP was discussed "some" to "quite a bit." However, interviewees from only one district--Koskitown--noted such a discussion.

Table 15.--Results of analysis of variance for Item 6b: MEAP has been discussed at teacher meetings.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
MEAP discussed at teacher meetings	9	132.724	14.747	4.32	.0001*
Error	84	286.978	3.416		
Corrected total	93	419.702			

*Significant at the .05 alpha level.

c. MEAP test results have been discussed at school board meetings.

This type of communication received a mean rating of 4.472 from respondents. This means that, on the average, MEAP test results had been discussed "some" but leaning toward "quite a bit" at school board meetings. The number of responses to this question was 87 instead of 93. In addition, significant variability in responses among districts was present. Data for this item are presented in Table 16.

Table 16.--Results of analysis of variance for Item 6c: MEAP results have been discussed at school board meetings.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
MEAP results discussed at board meetings	9	80.606	8.956	2.52	.014*
Error	78	277.348	3.556		
Corrected total	87	357.955			

*Significant at the .05 alpha level.

In the interviews, only staff members from Aholathi and Koskitown referred to the sharing of MEAP results at school board meetings. In both cases, the reporting of the incident was in the form of a complaint, such as "They came down hard on the elementary principal because of a smaller percentage of students in the top quartile this year. Can you imagine that?" The other comment was directed to "who" was reporting the results. The implication was that it should have been the building principal rather than the superintendent communicating test results as principals are involved in the administration and interpretation of test results.

- d. Inservice activities have been provided to help incorporate MEAP objectives into instructional programs.

The mean rating for this item was 2.691, indicating that, on the average, educators in small, rural schools had had "very little" help in the form of inservice activities to incorporate MEAP

objectives into their instructional programs. Districts varied significantly in responses to this item (see Table 17).

Table 17.--Results of analysis of variance for Item 6d: Inservice activities have been provided to help incorporate MEAP objectives into instructional programs.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Inservice activities	9	223.898	24.878	6.04	.0001*
Error	84	346.156	4.121		
Corrected total	93	570.053			

*Significant at the .05 alpha level.

Using the Scheffe test, two significant contrasts were found. District 10 differed significantly from all others, with a mean of 6.5 versus 2.2678. The educators in District 10 indicated that quite a few to extensive inservice activities had been provided to help incorporate MEAP objectives into instructional programs. Respondents from other districts indicated fewer inservice activities of this type had been provided.

Aholanti, Makivil, and Vilmi differed significantly from the remaining districts, with mean ratings of 1.0843 versus 3.3797. Educators from these three districts indicated that very few inservice activities had been provided to help incorporate MEAP objectives into the local curriculum. In interviews, teachers in

these districts were either restricted or did not care to travel to inservice sessions geographically removed from their districts. This restriction might have affected the access to information on such inservices or the perceived availability of such inservices. Respondents from other districts in this study perceived a greater number of inservice activities directed to this purpose.

Both administrators and teachers in Aholahiti, Koskitown, and Vilmi indicated that inservice activity that year had been targeted to the new definition of reading and effective schools. Each of the three districts named sent a delegate to the Department of Education/Midcentral Rural Education Laboratory series. These delegates then presented inservices to teachers in their respective districts aimed at familiarizing teachers with the new reading definition and providing strategies to teach reading in accordance with this new definition. Such a delegate was not mentioned in Makivil; teachers from that district attended few conferences.

Responses from Aholahiti and Vilmi educators were enigmatic. A delegate from these districts had attended approximately nine conferences. An equal number of modules was available for presentation in the local district.

- e. Our staff shares information regarding teaching strategies used to increase student test scores on MEAP.

Rural educators gave this item a mean rating of 2.872. This means that, on the average, staff members did "very little" to "some" sharing of information regarding teaching strategies used to

increase students' test scores. Districts also varied statistically in this regard (see Table 18).

Table 18.--Results of analysis of variance for Item 6e: Our staff shares information regarding teaching strategies used to increase student test scores on MEAP.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Sharing of strategies	9	133.129	14.792	4.29	.0001*
Error	84	289.339	3.445		
Corrected total	93	422.468			

*Significant at the .05 alpha level.

The Scheffe test indicated at least two significant contrasts. Makivil, Vilmi, and District 8 differed significantly from the other school districts in the sample; the respective mean ratings were 3.3468 and 1.3701. These three districts did "some" sharing of strategies to increase test scores, whereas the other districts tended to do less, actually "very little" sharing of strategies. At the other extreme, District 10 shared "quite a bit," and other districts shared "some" to "very little." The respective mean ratings were 5.60 and 2.4375. Variations were verified through the interviews.

In Aholahiti, several teachers indicated that they did not teach the test objectives. Thus, one might deduce that they probably did

not share strategies to improve test results. Said one teacher, "I don't do it [teach test objectives], and I administer it! I want to find out their true level [of functioning]." Another teacher said, "I don't prepare _____ graders for the _____ grade test. Is that good?"

Koskitown respondents presented an entirely different picture with regard to teaching test objectives and sharing teaching strategies. Said the superintendent, "Teachers are not reluctant to teach the test objectives. . . . They are concerned about the way they present the material." This was backed up by a statement made by the principal: "We do an item-by-item analysis, trying to discover ways to improve our scores."

Interviewees from Makivil were relatively silent on the sharing of teaching strategies, as were those from Vilmi. This correlated with their survey responses.

Summary of Communication of MEAP Results

Principals in small, rural schools provided "some" leadership in the interpretation of MEAP results. When counselors were present, the interpretation of results was often delegated to them. Respondents from different districts varied in their responses to this question.

MEAP was discussed "some" at teacher meetings; some districts did this more intensely than others. In Koskitown, an entire teachers' meeting was devoted to such a discussion.

MEAP test results were discussed "some" at school board meetings, with variability present among districts. These discussions were reported in Aholahti and Koskitown.

"Some" inservice activities had been provided to help incorporate MEAP objectives into instructional programs. Vilmi, Aholahti, and Koskitown respondents indicated that they were involved in reading activities. Respondents from different districts varied significantly in their responses to this item.

"Very little" to "some" sharing of information regarding teaching strategies used to increase student test scores occurred in small, rural districts in the Upper Peninsula. However, respondents from different districts varied significantly in their responses to this item.

Questionnaire Item 7

The intention of Item 7 was to investigate the type and amount of change that had occurred as a consequence of MEAP test results. Small, rural school educators responded to seven possible changes. A discussion of the type and number of changes is provided in the following paragraphs.

If variability was present, the Scheffe test was used to identify significant contrasts. A discussion of these identified contrasts is provided. However, other contrasts might be present and significant but remain unidentified.

a. Course offerings.

The mean rating for this change was 2.478, indicating that, on the average, rural educators perceived "very little" change in the course offerings based on MEAP test results. Further evidence of this was provided by the interviewees.

Aholahti was described as a "bare bones" district. It was providing the essentials and could afford little more. Koskitown educators described their curriculum as "very basic." If a course were to be added, a foreign language would be considered. Educators in Makivil and Vilmi were silent on this issue.

b. Course content.

Rural educators gave this change a mean rating of 3.413. This means that, on the average, educators indicated that "some" change had occurred in course content based on MEAP test results. The variation among districts was statistically significant (see Table 19).

Table 19.--Results of analysis of variance for Item 7b: Changes in course content based on MEAP results.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Course content change	9	83.852	9.317	4.23	.0002*
Error	82	180.452	2.201		
Corrected total	91	264.304			

*Significant at the .05 alpha level.

One significant contrast was found using the Scheffe test. Makivil varied significantly from the remainder of the districts in the amount of course-content change that had occurred because of MEAP test results. Makivil educators indicated "very little" change had occurred, whereas those from other districts tended to indicate that "some" change had occurred. Mean ratings were 1.0 and 3.6640.

Respondents from the small, rural districts indicated variability in the interviews also. In Aholanti, teachers in the past had met and discussed weaknesses in their curricula based on MEAP results. Teachers tended to focus the discussion on math. One teacher said, "Our lowest part is decimals and fractions. We stress them in fifth and sixth grade. Now the teachers must get to them." However, one of these targeted teachers indicated, "I hit hard on fractions and decimals but not in the same manner as in the test. I use various texts and give them a different presentation from that of the selected math series. Kids don't adapt well to variability in format." Administrators indicated that discussions of test results were perceived to alter instruction, but they did not check on it. The counselor indicated that the likelihood that teachers altered instruction and content was better than 50%.

One teacher summed up the majority of educators in Aholanti in the following manner: "I don't believe in teaching the test objectives, but I would teach general objectives. Teachers have decision-making power; I can do it and I enjoy doing it."

In Koskitown, teachers made a conscious effort to teach the test objectives.

Teachers in Vilmi tended to align with this teacher's comment: "I don't believe in teaching toward the test. What is covered in the MEAP test, I am covering in my classes. The curriculum does meet the test although all the materials are not there prior to the test. Some of the material comes later."

Educators in Makivil were silent on this issue during interviews. Course content was determined by the textbook and experience; MEAP really did not have a role in course-content change.

c. Teaching methods.

The mean rating for this change was 3.366, indicating that, on the average, rural educators changed teaching methods "some" based on MEAP test results.

Koskitown was the only district whose respondents alluded to the sharing of strategies based on test results. In this district, suggestions were implemented as a team; individual teachers were not requested to make changes.

d. Instructional methods.

For this change, rural administrators and teachers gave a mean rating of 3.366, indicating that, on the average, "some" changes were made in instructional methods based on MEAP test results.

e. Preparation for tests.

The mean rating for this item was 3.283. This means that respondents, on the average, changed "some" with reference to preparation for tests.

Aholahti's administrators indicated that they did nothing special to gear students up for tests. They had sample tests, as for the ACT, and announced a seminar held in an adjacent community to prepare for the ACT.

Koskitown's preparation for MEAP tests was on-going due to the importance of the "pre-MEAP" tests and the use of test results to group students by ability. Students in the elementary grades knew that these tests determined who would be in the high-ability groups. In Koskitown, the importance of these tests was also indicated by having the high school principal administer MEAP tests. The belief that "testing is the best instrument" to measure effectiveness tended to permeate the core of Koskitown. However, as many as would agree that tests are the best measure, an equal number of teachers would also agree that they are not the only measure. If it is important, continuous preparation is taking place.

Vilmi did some test preparation, but its actions were not organized. One classroom teacher had a segment "where I go over test procedures and do practice test types of things. There are no grades taken and therefore no pressure. There are students who do feel a lot of pressure on test day. . . . I make it clear to students that the test is not going to make or break them." Another

classroom teacher did an interesting exercise with his students. He described it as follows:

I gave students an answer sheet without a test booklet. I indicated that they had 15 minutes to do 93 questions. The top three students would be rewarded. After 15 minutes, we corrected the test. I would tell them, "Let's correct the sample questions first--they are the easiest." I asked how many got these right. Then I went into a spiel about, "You didn't get them right, how come? etc." At the end of this exercise, the highest scorer was 25 correct out of 93. The intent of the exercise was that guessing doesn't pay.

This teacher indicated that marginal students took a little more time with test completion after the exercise.

f. Battery of tests.

Rural educators, on the average, changed the battery of tests "very little," as indicated by the mean rating of 2.429.

In Aholahiti, college and armed services entrance exams were administered in addition to the MEAP. Recently the superintendent discovered that some special education students could be excluded from taking the MEAP tests. This had not been practiced in Aholahiti and was now being considered "so that we can look better. We do all right, but we can do much better if we exclude them. One kid could represent 4% of our population." Thus, although the battery of tests had not altered, the population of students tested might change as a means to improve overall MEAP scores. Change might occur in the future as a number of teachers in Aholahiti were hoping for more testing, better interpretations, and greater use of test results.

In Koskitown, the California Achievement Test, MEAP, and tests associated with college or armed services entrance were given. MEAP had not altered the district's test battery. The superintendent had asked his principals, "Why do we give the California Achievement Test?" Despite the considerable expense, principals had indicated only that the test results confirmed teachers' perceptions and that the test was used as a gatekeeper for their compensatory education programs. However, the superintendent indicated a need to explore the curriculum-test link.

Educators in Makivil were silent on the issue of a test battery, although they did administer the Stanford Achievement Test and tests in the college and armed services entrance battery. Test results just were not the focus of this district's energy.

Vilmi administered the SRA, the MEAP, and a battery of tests intended for entrance into college or the armed services. Educators tended to think that testing was important so that "students can be compared and to compare them to a statewide average or national average. Students as well as teachers want to know." In Vilmi, the method of how tests were selected was not well documented; it was assumed that tests were selected on the basis of a recommendation made by a compensatory education consultant. The SRA was recommended because it allowed more students into this "remedial" program. Of the MEAP and the SRA, the latter was of greater value.

Respondents from Aholathi indicated that, if the MEAP were not mandated, they would not have the money to continue it. "It is a good idea," said the superintendent. "We would need another test

similar to it. Yes, we'd give another test, probably the Stanford." The superintendent in Vilmi also indicated that the MEAP would not be given if it were not mandated. As in Aholanti, it would be replaced with something. Both superintendents indicated that the MEAP had become a part of the school system; it was expected. Thus, it must be replaced.

g. Frequency of curriculum revision.

The mean rating for this change was 2.667, indicating that, on the average, administrators and teachers altered the frequency of curriculum revision "very little." However, districts varied significantly in their collective responses to this item, as shown in Table 20.

Table 20.--Results of analysis of variance for Item 7g: The change in the frequency of curriculum revision based on MEAP results.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Frequency of curriculum revision	9	76.232	8.470	2.58	.0113*
Error	83	272.434	3.282		
Corrected total	92	348.667			

*Significant at the .05 alpha level.

Two significant contrasts were identified with the Scheffe test. Makivil and Aholanti differed significantly from Koskitown and District 10, with mean ratings of 1.9523 and 3.9045. The MEAP played a minuscule role in curriculum revision in these districts, whereas the test played a larger role in Koskitown and District 10. It had increased the frequency of curriculum revision "some."

Makivil, Aholanti, and District 3 differed significantly from Koskitown and District 10. MEAP had not affected the frequency of curriculum change in Makivil, Aholanti, and District 3 as it had in Koskitown and District 10. The mean ratings were 1.5879 and 3.9045.

Educators in these small, rural school districts did not comment extensively on curriculum revision. However, it was going on, as indicated by the fact that all of the districts in which educators were interviewed were struggling to incorporate the new reading definition into their reading programs. Koskitown educators indicated that they did not have all the curricula in place; however, they continually strove to reach that goal. Roughly 50% of this sample could revise the existing curriculum. The remaining 50% needed to work on developing curricula, as indicated by responses to Item 4.

As in Item 6d, the responses of Aholanti and Vilmi educators remain a puzzle because the survey and interview responses were incongruent. Personnel from both districts had attended a rather large number of inservice sessions on the new reading definition for the MEAP reading test.

Questionnaire Item 8

Item 8 comprised four parts. The first item addressed the familiarity of teachers with MEAP test questions. The next two items measured the ownership of MEAP test scores by students and their parents. The last item assessed the concept of remediation for MEAP objectives not mastered by individual students. These four items are discussed in the following paragraphs. If variability was present, the Scheffe test was used to identify significant contrasts. The contrast found will be discussed; however, additional contrasts might be present and significant but remain unidentified.

a. Teachers are familiar with MEAP test questions.

The mean rating for this item was 4.128, indicating that, on the average, rural educators were somewhat familiar with MEAP test questions. However, respondents from different districts varied significantly in their responses to this item (see Table 21).

Table 21.--Results of analysis of variance for Item 8a: Teachers are familiar with MEAP test questions.

Source of Variation	df	Sum of Squares	Mean Square	F-Value	p-Value
Teachers are familiar with test questions	9	69.037	7.671	2.24	.0268*
Error	84	287.431	3.422		
Corrected total	93	356.468			

*Significant at the .05 alpha level.

One significant contrast was found. Aholanti, Makivil, and Vilmi differed significantly from Koskitown and District 10, with mean ratings of 3.0485 and 5.4227. Educators from Koskitown and District 10 indicated that they were "quite" familiar with MEAP test questions; Aholanti, Makivil, and Vilmi teachers had "some" familiarity with these questions.

As the superintendent in Koskitown reported, "Teachers have studied the manner in which the material is tested." They were concerned about the question format and the way they presented the material. This was the only direct comment made on this issue, but it supports the findings from the survey.

b. Students know their MEAP scores in our district.

Rural educators gave this item a mean rating of 5.527. This means that, on the average, quite a few students knew their MEAP scores. Students in Koskitown and Vilmi had the greatest probability of knowing their scores. Tests were important in Koskitown. In Vilmi, "students want to know their score," said one teacher.

c. Parents know their student's MEAP score.

Administrators and teachers in small, rural school districts gave this item a mean rating of 5.693, indicating that, on the average, quite a few parents knew their students' MEAP scores. Interviewing did not produce any comments on this subject.

- d. Remedial classes are designed to teach students the mastery of specific MEAP objectives.

The mean rating for this item was 2.233. This indicates that very few remedial classes were designed to teach students the mastery of specific MEAP objectives.

Compensatory education teachers indicated that the MEAP test results could be included as one criterion on their needs assessment. However, these results were used only for screening; they were not used to plan an instructional program and certainly were not used as part of the exit criteria. Said one compensatory education teacher, "I am hoping to use it [MEAP] more."

In one district, Koskitown, students received additional instruction, based on MEAP test results. If performance was not satisfactory, students were enrolled in a "remedial MEAP class" and remained in that class until they passed the test. This occurred with seventh and tenth graders in Koskitown. Elementary students were required to pass "pre-MEAP" tests before being promoted to seventh grade. These tests were administered in fourth, fifth, and sixth grades. Over the years, the "pre-MEAP" tests have remained, but the "remedial MEAP class" has been discontinued. The superintendent indicated that this combination, remedial classes and pre-MEAP tests, provided the motivation for students and structure for teachers.

Four Vignettes

The researcher's intention in this section is to provide a description of four small, rural districts in the Upper Peninsula of

Michigan, each of which was surveyed in this study. All districts had enrollments of 500 or fewer students.

The four districts were geographically removed from the state capital and the Department of Education by the Straits of Mackinaw; a bridge connects the peninsulas. One district was approximately 500 miles from Lansing, whereas another was just 250 miles from Lansing. The Michigan Department of Education had acknowledged the problems associated with distance and had provided the Upper Peninsula Outreach Office in Marquette.

All districts classified their communities as isolated and depressed. Gjeltén (1982) associated the problems of transportation, commerce, and cultural activities with the isolated, rural community. In addition, economic insecurity abounds, outmigration is high, and the local economy is underdeveloped. Superintendents were in agreement with such a description; however, it seemed to vary in terms of amount of deprivation as well as recency.

All districts in which interviews were conducted had an average 1988 MEAP ranking of 75 or better (Public Sector Report, 1989). This means that, across all reading and math tests, the average ranking was 75--a satisfactory ranking on the MEAP test. Three of the districts had a ranking in the 80s.

Districts were similar in some respects but dissimilar in others. These districts varied on 11 items on the survey instrument, indicating "between"-district differences. Interviews usually verified such differences. Additionally, variability was

present "within" the districts, as witnessed by the lack of uniform responses to questions both on the survey instrument and in the interview. Responses within each district varied on all survey questions except questions 2 and 4. Variability that is statistically significant is shown in Appendix Table D.4. However, the researcher advises extreme caution when interpreting within-district differences because of the small number of responses in each district. Thus, statistically significant differences existed between and within small, rural districts in this study.

A description of the four districts in which interviews were conducted follows.

Aholahti

Aholahti was nestled in rocky bluffs adjacent to one of the Great Lakes. A federal highway passed through the town of approximately 2,500 inhabitants. Businesses greeted one at the east gate as they bade one farewell at the west gate. Aholahti was one of several small communities within a ten-mile radius.

The economy of Aholahti was quite dependent on the service industries. The community relied on tourism during all four seasons, which included a growing snowmobiling trade, a stable charter boat and fishing trade, and a developing hiking and camping business. Aholahti had one primary manufacturing industry, which seemed to have stabilized with new management and aggressive selling. In the recent past, it had experienced years of financial loss and bankruptcy. Other smaller businesses were present, which

included a sawmill and several gift shops. Several service stations and restaurants peppered the edges of the highway.

Aholahti was located in a county where the median age was 39. Its aging population comprised 21% of the county citizens. Approximately 50% of the households included someone who was 60 or older. Aholahti's superintendent perceived that there were two types of new residents in the community: alumnae who had returned to retire, and one-parent families seeking public assistance.

The educational level in Aholahti varied, but a majority of residents had a high school diploma. However, 34% of the populace had not received a high school diploma. Ten percent of the county's residents were college graduates.

Aholahti was unique in its prevalent and documented pride in educational successes. In recent years, doctors practicing at the University of Michigan, Mayo Clinic, and elsewhere had been asked to return to address the graduating classes. The high school principal, himself a graduate of Aholahti, said, "More doctors, engineers, attorneys, and educators have come out of this school percentage-wise than any school I know of." These professionals made the pilgrimage to Aholahti not only for graduation, but for reunions on the Fourth of July, Labor Day, and so on. The school had educated its clientele to leave the area; yet there remained the yearning to return to the community whenever possible.

The median household income for the county in which Aholahti was located was \$10,972. The state's median income was \$28,956 at

the time of the study. The superintendent also believed the community fell into the low-middle socioeconomic category.

The Aholahti district was made up of an elementary and a combined junior high/high school. The buildings were erected at about the turn of the century; one building's structure was partially wooden and was in great need of repair.

Aholahti had a superintendent, a high school principal, and a principal/teacher as its administrative staff. All were indigenous to the area, as were many of the teachers. As indicated by this administrative staff, all teachers were instructing within their areas of certification. Thirty-six teachers provided instruction to Aholahti's 500 students.

The district's average MEAP score was over 80. Aholahti educators indicated that they really did not do much with MEAP test scores. In the past, teachers had looked at the math profile, found weaknesses in decimals and fractions, and stressed these concepts in the fifth and sixth grades. These teachers, as a whole, tended to be opposed to test objectives. Tests were to yield a "true" score. Teachers did not feel pressured to incorporate MEAP objectives into daily instruction, although most acknowledged that they covered the content areas of the MEAP.

MEAP test results typically were communicated to the school board and appropriate committees by the superintendent. The superintendent also reported the results to the teachers.

Aholahti was found to vary significantly from other districts on the following items:

1. MEAP affected the frequency of curriculum revision "little."
2. MEAP was used "very little" to identify staff-development needs.
3. Very few inservice activities were provided to help incorporate MEAP objectives into instructional programs.
4. Very few teachers were familiar with MEAP test questions.

The researcher speculates that these dimensions vary due to a lack of leadership in curricular areas. The superintendent appears to be providing the only direction, which is limited; without direction, teachers are making decisions independently of each other. Additionally, these teachers are not communicating decisions with fellow staff members.

An additional speculation is that the long tenures of both the superintendent and the high school principal have been attested by student successes. Under their leadership, students have successfully completed college, vocational school, or the armed services and have entered and achieved recognition in their chosen professions. Very few community members would challenge the success of the students under their leadership.

One further observation seems appropriate. Budgetary practices with respect to procurement of textbooks are suspect. In the past, only one or two grades were able to purchase a textbook. Greater communication and coordination might occur if a textbook-adoption

process were endorsed and if textbooks were purchased less frequently but for all elementary grades in a particular subject matter.

Despite the lack of usage of MEAP test results, tests were important in Aholathi. District personnel were contemplating the addition of the Stanford Achievement Test to their test battery in the elementary school. The high school would continue to give the typical battery of tests for students bound for college or the armed services.

Educators in Aholathi indicated that conferences were the best method of infusing their district with common ideas. The intermediate school district to which they belonged had provided that service well. However, teachers would have liked some of the "presenters" to provide inservice in their district.

Overall, Aholathi students had scored very well on tests in the past. There was the perception that they would continue to score well despite the decline in the caliber of students arriving at the schools. The curriculum-text-test link had not been explored, nor was it of particular interest in this district at the time of the study. The selection of textbooks was uncoordinated, test selection was based on the recommendations of others and was unrelated to textbook selection, and the curriculum was not defined in the elementary school. The curriculum-text-test link was not perceived to be the gray cloud on the horizon; keeping the schools open and

operating consumed the time and energy of educators in this school district.

Koskitown

Koskitown was located on a major highway traversing the Upper Peninsula of Michigan. When approaching Koskitown, the traveler is treated with spectacular views of inland lakes, rocky cliffs, and one of the Great Lakes. Koskitown had one traffic light, with small businesses clustered around this intersection.

The district was roughly 300 miles square and had to deal with a rather unusual problem: There was virtually no land available to buy for development purposes. Its shorelines were privately owned or state or federally owned. As much as one would expect condominium development on the available beaches, such development was not found.

Koskitown was located in a county where approximately 10,000 people resided. The median age of residents was 32.5. Approximately 17% of the residents were 65 or older. The superintendent of Koskitown indicated that new arrivals tended to be retirees; some were well educated, whereas others had retired from the assembly line.

The median level of education for Koskitown was 12.3 years. Approximately 9% of the residents were college graduates; 14.2% of the residents in Michigan were college graduates at the time of this study. The current dropout rate ranged from 0% to 6%; most students dropped out of school due to pregnancy.

Koskitown measured success by the number of students attending college, the high attendance record, and a low dropout rate. But, unlike other districts in which interviews were conducted, Koskitown "looked a lot" at the MEAP, the California Achievement Test, and the ACT. Teachers indicated that "testing is the best instrument [we have] but not the only one."

The estimated median household income for the county in which Koskitown was located was \$17,280. The median household income for Michigan was \$28,956. Household incomes could be generated from the mining industry, governmental service, and tourism. There was little employment for young people.

Koskitown's school was one facility. The older school had nearly been engulfed by additions. Approximately 25 teachers instructed 350 students. The atmosphere of this school was businesslike but also friendly. Tests helped measure success in this school, as one teacher stated: "Students don't see them [tests] as important; teachers do." In fact, some teachers thought that administrators used test results to evaluate them.

Because test results were valued, teachers in this district spent time culling the results, developing strategies, and accepting responsibility for instruction of test objectives. The high school principal administered the test to junior high and high school students. Elementary students were motivated to do well on tests because these measures were used to sort classes into high-ability and moderate-ability groups. Students also took a "pre-MEAP" test before leaving sixth grade.

Koskitown was actively developing and upgrading curricula. The superintendent indicated that they were striving to develop the curriculum first, and then select the textbook. Koskitown personnel had not looked systematically or carefully at the curriculum-textbook-test link, but they were aware that it must be done sometime in the future. They also realized that teachers were not as familiar with test interpretation as they should be. Many in Koskitown wanted to do more with test results. The superintendent wanted test results to have greater consequences.

One would be remiss to allow the reader to believe that all educators in Koskitown were enthralled with the test battery. Some members "avoided them on purpose." This avoidance tended to be selective to particular tests as this same educator indicated that the "new reading test is just great!"

In contrast to other districts in this survey, Koskitown respondents believed that MEAP increased the frequency of curricular revisions. Teachers were quite familiar with test questions, and students were placed in remedial programs based on their MEAP test results.

The interviews conducted in this district clearly indicated that administrators were actively involved in the administration and interpretation of test results. It is suspected that the superintendent sought additional information on testing and provided the leadership for administrators and other staff members. Thus, MEAP results were discussed at teachers' meetings, and individual

planning sessions were conducted for teachers in the third, sixth, and ninth grades to prepare students for the test.

Administrators in Koskitown have asked for a process to align the curriculum, instruction, and the test. Given the process and access to answers to technical questions, this researcher believes that these administrators can lead to alignment.

This district varies from others in that additional education is prized, encouraged, and financed. Staff members attend conferences, inservices, and graduate courses. Staff members are required to discuss the contents of their educational pursuits at staff meetings.

Finally, school board members are sensitive to variations in test results. They expect improvement and question principals when MEAP results do not reflect improvement.

Thus, Koskitown varies from other interviewed districts in human resource and process dimensions.

Rural educators in Koskitown were satisfied with the inservice activities provided by their local intermediate school district. One member indicated that "we don't need a motivational speaker, we need an opportunity to share ideas with teachers in similar grades or situations." The superintendent indicated that assistance in curricular issues would be helpful. The principal indicated that having "experts" visit the district would be beneficial in school improvement. Thus, the focus of educators in Koskitown seemed to be the revision and refinement of their curricula.

Makivil

A state highway was traveled to arrive at Makivil. The ascent to the village was picturesque, encompassing a large inland lake and surrounding forests. This was a town without a traffic light but with a majority of businesses clustered on main street. The population was approximately 300. Neighboring villages were about 30 miles away.

The district encompassed approximately 250 square miles. The students resided close to the community so that, at the time of the study, they were transported to the school by public transportation.

This district, much like Koskitown, had very little land available to buy. A majority of its land was in commercial forest, federal reserve, or a sportsmen's club. Less than 15% of Makivil's land was on the tax rolls.

Makivil was located in a county of approximately 8,000 people. The median age in this county was 31.7; the median age in the state was 28.8. Approximately 19% of this county's populace was over age 65. The superintendent described this retired population as highly educated.

The median level of education in Makivil was 12.3 years. Approximately 10% of the residents were college graduates; the state average was 14.2%.

The estimated median household income in the country was \$21,168. The state median household income was \$28,956. Makivil was the largest employer in the village. Tourism was on the rise and was currently limited by the number of "year around" restaurants

and hotels available. The logging industry also played a role in the village. People worked "two or three jobs here because they want to be here," said the superintendent.

Makivil had one well-kept, two-story building. Approximately ten teachers delivered instruction to 100 students. The superintendent tended to be the administrator, although one teacher was recognized as the principal: "[I am] principal by default, I do it because it needs to be done." Teachers had nine preparations; thus, several were teaching out of their areas of certification. Despite the assignment, it was not difficult to attract or retain teachers. This was a faculty who worked well together; exchanging of favors was common but was still "very much appreciated."

This district had experimented with the electronic blackboard, linking itself with a much larger district. Difficulties with a common calendar, class hours, and teacher agreements caused this program to be disbanded. However, Makivil was once again looking at some long-distance learning.

Makivil measured its success by the low dropout rate, the number of students going to college, and test scores. Makivil educated its students to leave; students usually selected the service or college.

Makivil educators tended to agree that tests provided useful information about school effectiveness; they used MEAP in conjunction with other tests. They did not report any systematic procedure to review the test results or develop strategies. In

fact, most indicated that instruction was based on a combination of textbook material and personal experience.

Teachers in Makivil indicated that the curriculum was in place but "not objective based." The curriculum-text-test link was yet to be explored in Makivil. The math-science teacher administered a battery of tests that he found useful, and he altered instruction based on the test results. However, most "don't teach for the test; . . . they teach for the whole child to get him through life." Observation and student feedback played a very important role in Makivil as "I have them for three hours."

Teachers in Makivil varied significantly from those in the other districts by indicating that MEAP had little influence on the following items:

1. Course content.
2. Frequency of curriculum revision.
3. The number of inservice activities provided to incorporate MEAP objectives into the local curriculum.
4. Use of MEAP to identify staff-development needs.
5. The sharing of strategies to increase student test scores.
6. Familiarity with test questions.
7. The principal providing leadership in test interpretation.
8. Discussion of MEAP at teacher meetings.

Based on interviews in Makivil, the researcher suspects that leadership is lacking with respect to curriculum. In this district, the superintendent, as sole administrator, must assume this responsibility. It appears that little direction is provided, with

each teacher making independent decisions and not communicating these decisions to others.

This district's smallness should enhance communication. However, it, along with Vilmi, communicated MEAP test results least at meetings and with teachers.

Makivil appears to lack the leadership and communication that are present in Koskitown.

Educators in Makivil felt alienated from their intermediate school district, not only by distance but by issues of concern. They had neither the desire nor the interest to attend workshops. Makivil teachers' energy was focused on a "fight for our life," primarily survival. However, students tended not to slip through the cracks with such intense observation as was found in Makivil.

Vilmi

Vilmi could be approached from the east and west by a federal highway. One of the town's two traffic lights marked this intersection. Several small businesses clustered around this intersection, and the remaining businesses were scattered throughout the town. The 1,000 residents lived along the highways, rivers, and shorelines of one of the Great Lakes. The nearest towns were approximately 30 long, desolate miles away.

Vilmi was located in a county that had approximately 20,000 residents. The median age of these residents was 39. Approximately 30% of the residents were 60 years or older. The school enrollees comprised 25% of the populace.

A majority of residents in the county in which Vilmi was located had a high school diploma. Approximately 10% of the residents were college graduates.

Like other rural districts, Vilmi measured its success by the number of students going to college. However, teachers further defined success as "being successful in the field they want to, be it the job market, be it vocational school; being able to, in some meaningful way to them, fit into the world in which they want to. In other words, being successful in themselves." Interviewees indicated that 100% of the graduating seniors would be attending college the following fall; the percentage attending college was usually about 80%.

The median household income was \$10,972. The median household income for the state was \$28,956. Household incomes in Vilmi were generated from governmental work, logging, tourism, and service industries.

Vilmi's school complex consisted of one building. The center of the structure was the old two-story school; additions had been made to the west and east sides. In this facility, approximately 15 teachers provided instruction to 125 students. Educators described the inhabitants of this building as one big family, working well together and caring for each other. Said one staff member, "It is a closeness that is hard to find in the schools of today."

Historically, this school district was known for its academic success story, being included in the effective schools research. It

had been able to rise above a low socioeconomic status to produce high test results. At the time of the study, Vilmi administered the SRA test to all its students. Other tests included in its battery were the MEAP, ACT, SAT, and ASVAB.

As the superintendent indicated, "tests do not play a major part" in the assessment of the district. Teachers tended to believe that tests were not taken seriously by the students. Teachers had reviewed the MEAP test results, altering the English course content to include instruction on prefixes. However, teachers in this district tended to agree that "I don't believe in teaching toward any test. I don't think that is valid." They tended to believe that "what is covered in the MEAP test . . . is covered in my classes. The curriculum does meet the test although the materials are not there prior to the MEAP test." Some instruction was presented on test-taking skills in Vilmi.

Overall, teachers in Vilmi believed tests were useful in that students could be compared to a statewide or national average. The curriculum-text-test link had not been discovered, although an effort had been made to align the new reading definition to a reading textbook. The superintendent had raised the concern as to how Vilmi would continue to show improvement when its students were already scoring very high. It was obvious, from the interviews, that this school wanted to be considered an "improving" school. Vilmi educators' time and energy could be directed to curriculum issues because the district was not plagued with the financial worries of most small, rural school districts.

Vilmi, along with Makivil, varied significantly from the other small, rural districts studied in the use of MEAP results in the following ways:

1. Inservice activities provided to incorporate MEAP objectives into the local curriculum.
2. Identification of staff-development needs.
3. The sharing of teaching strategies to increase student MEAP scores.
4. Familiarity with test questions.
5. The principal providing leadership in the interpretation of test results.
6. The communication of test results at teacher meetings.

These two districts indicated that they used results "very little," whereas other districts indicated that results on the aforementioned items were used "somewhat."

The researcher suspects that the emergent pattern for Vilmi and Makivil stems from the personal affiliations of staff members. "One big family" may communicate it best, but the issue of test usage probably is not viewed as a high-priority item. Additionally, if discussed, these exchanges are not formalized so as to be considered at teachers' meetings.

Second, the researcher suspects that neither administrator in Vilmi or Makivil feels comfortable in leading their staffs in the development or revision of curriculum.

Summary

Four vignettes were presented to acquaint the reader with these small, rural districts in the Upper Peninsula of Michigan. The districts were similar in that they all scored high on the state-mandated MEAP test, as well as on other tests. All were described as declining, rural districts. In these districts, students were educated to leave.

Despite the similarities among districts, differences were also present. Educators in Koskitown tended to review and implement strategies to raise test scores. Makivil had curricula in place. Aholahti had an inordinately high percentage of physicians, engineers, attorneys, and educators who were graduates of its high school. Historically, Vilmi students had scored high on tests and continued to do so. A review of the surveys and interviews indicated that similarities as well as differences existed in small, rural districts of Michigan's Upper Peninsula.

Chapter Summary

Percentages were presented in the first section to indicate the educators' beliefs about the usefulness of standardized tests to measure effectiveness, the use of MEAP in conjunction with other tests, and the presence of curriculum guides. These percentages were presented as a district mean, as well as a collective mean of the districts sampled.

In the second section of this chapter, means were presented for the sampled population. Significant differences in district means

were found regarding the use of MEAP tests to determine placement of students in "remedial" programs and to identify staff-development needs. Districts also varied on the dimensions of principal as the interpreter of MEAP results, discussion of MEAP at teacher and school board meetings, inservice programs targeted at the incorporation of MEAP objectives into the local program, the sharing of strategies to raise MEAP scores, the altering of course content based on MEAP scores, the frequency of curriculum revision, and teachers' familiarity with MEAP questions.

In the last section, the four districts in which interviews were conducted were presented as vignettes. The districts were similar, yet in some ways they were quite different.

These districts varied significantly on 11 items in the survey instrument. Variability was found between and within districts during the interview process. In the domains of MEAP as a part of the testing program, curricular decisions, course content decisions, and communication, variability was present. Variability was found when responses indicated "somewhat" as well as "very little" resulted from MEAP test results. Also, significant variability was found in "low" effort variables, as in teacher familiarity with test questions, as well as in "high" effort variables, as in curriculum revision.

A pattern that seems to emerge is that the two smallest districts tend to vary significantly from larger districts. The researcher suspects that these differences are due to human resource

and process deficits. Specifically, leadership in curricular issues and communication of test results seem to be lacking.

In Chapter V, a summary of the major findings, conclusions based on those findings, and recommendations for further study are presented.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The federal- and state-level initiatives of the 1960s and the 1970s affected local educational programs. The primary concern of these initiatives was to provide equal educational opportunities to all children. In the late 1970s and early 1980s, the focus of education began to shift to accountability, but equity remained in the foreground. States assumed greater responsibility in the education of children; testing was selected as the least intrusive means of monitoring and controlling. As a result, mandatory state assessment programs appeared. These programs became prevalent, with all states considering an assessment program of some type. Consequently, a majority of states have mandated assessment programs, which strive to deal with the accountability and equity issues.

The Michigan Educational Assessment Program (MEAP) was Michigan's means of assessing and addressing the issues of equity and accountability. MEAP legislation was enacted in 1969 amid and followed by considerable controversy. The MEAP test was designed to assess educational needs: "to help children learn" and "to inform the public about its schools," as the MEAP handbook (1985) states.

In helping children learn, MEAP results were to be used in analyzing the strengths and weaknesses of local curricula. Local curricula were to contain the "essential objectives." The administration of the MEAP test is mandated; the communication or use of test results is not.

Publications by the Michigan Department of Education, the MEAP handbook in particular, provided detailed instructions for the use of MEAP test results in curricular assessment and the communication of MEAP test scores. The Department of Education provided the structure and encouragement; the actual assessment and communication were left to local school districts.

The communication of MEAP test results, especially the publication of the results in the newspaper, is one of the strongest forces to cause teachers to alter behavior. The Department of Education, in making public these results, is using one of the strongest forces available to cause local districts, through the teachers, to align with the test. But the Department is not limited to publication of test results; it recommends curricular revision, providing districts with "essential objectives" to be included in such curricula. Thus, the Department of Education couples the two strongest forces together to cause teachers to change. What remains is the response of districts to these forces.

Studies have been conducted to assess the communication and use of test results. In general, teachers did not perceive that test information altered instruction or curricular content, but that tests tended to confirm their perceptions of individual students'

achievement. However, as tests started to match the curriculum content more closely, as in criterion-referenced tests, the results of such tests became more meaningful. Teachers acknowledged the usefulness of tests for a common focus but resented the "grading" of schools on the basis of tests alone.

Of particular interest to the researcher were studies of the MEAP. MEAP was accepted and used, as indicated in the studies by Aquino (1975), Steele (1976), Jencka (1990), Thelander (1979), and Bushaw (1988). Their research samples included teachers, principals, superintendents, and school board members. The populations sampled tended to be the state of Michigan, but in Thelander's study the population was limited to a county. To summarize these studies, MEAP was used in developing the districts' goals and curriculum, determining the strengths and weaknesses of the present curriculum, diagnosing student needs, and reporting to the public. Teachers reported that they had altered their instructional methods to improve the students' test scores. However, the formalization of policy statements and policy actions related to MEAP did not, in itself, lead to improved MEAP scores. Teachers remained opposed to the use of MEAP test results for comparisons and for district evaluations.

Of the studies previously conducted, none focused on small, rural schools. Rural communities are complex and differ from each other. They are defined in terms of sparsity of population and distance from urbanization. Small, rural school districts are thus

defined in the same manner: small student enrollment and degree of isolation. Gjelten (1982) suggested an additional factor--the community in which they are located.

Although difficult to define, small, rural schools remain subject to the mandates of the state and the pressures associated with such mandates. They have problems associated with being small and rural--for example, attracting and retaining especially knowledgeable administrators--which must be surmounted in order to compete with their urban and suburban counterparts. To compete effectively is to survive.

Of interest in this study was the degree to which the state-mandated assessment test had been assimilated into the curriculum of small, rural school districts. Second, of interest were the ways in which small, rural school districts were intentionally trying to raise students' scores on the MEAP in an effort to compete more effectively--thus, to survive.

Purpose of the Research

The researcher's purpose in this study was to investigate the effect of MEAP on the curriculum in small, rural school districts. MEAP was examined in conjunction with the curriculum components of guides, instructional content, instructional strategies, the test battery, and communication of test results. The research questions investigated were:

1. Are there significant differences among districts with respect to the use of MEAP results?

2. What practices, if any, are intentionally used by small, rural schools in response to the MEAP?

Methodology

To determine the effect of MEAP on small, rural school districts, the researcher surveyed and interviewed educators from a sample of small, rural districts with enrollments of 500 or fewer students. Superintendents, principals, and third-, fourth-, sixth-, seventh-, ninth-, and tenth-grade teachers in ten Upper Peninsula school districts received the survey, which contained 33 forced-choice items. Interviews were conducted in four of the ten districts surveyed. These four districts were selected on the basis of isolation, student enrollment, type of community, and academic history. Analysis of data was used to (a) provide answers to the research questions, (b) indicate levels of MEAP usage, and (c) provide a description of small, rural schools.

Findings

The following is a summary of the general levels of MEAP usage, significant findings concerning the two research questions, and a profile of four small, rural districts in the Upper Peninsula of Michigan.

General levels. Educators from the small, rural school districts believed that standardized tests provided important information when assessing their districts' effectiveness. A majority of districts used MEAP results in conjunction with other standardized tests to aid in decision making. The selection of a

test to measure items of the local curriculum not found on the MEAP brought forth mixed results; about half of the districts did so, and half did not.

The presence of curricula in reading and math was found in about half of the districts. Of those having curricula, some of the "essential" objectives were included in them, textbooks were "somewhat" selected on the basis of these guides, and the guides "somewhat" influenced classroom instruction.

MEAP test results were used "some" to determine students' general achievement level, to inform the community, to determine strengths and weaknesses of the math and reading curricula, to determine instructional priorities, to determine placement in remedial programs, and to determine the need for new programs. These test results were used "very little" to analyze teacher performance, to identify staff-development needs, or to predict students' future academic success.

Principals provided "some" leadership in the interpretation of MEAP results; the results were discussed "some" at school board meetings and teachers' meetings. Staffs shared "some" information about teaching strategies used to increase MEAP scores. However, "very little" had been done to incorporate MEAP objectives into instructional programs.

The following had changed "some," based on test results: course content, teaching methods, instructional methods, and preparation for tests. "Very little" had changed in course

offerings, the battery of tests, or the frequency of curriculum revision, based on the MEAP test results.

Teachers were somewhat familiar with the test questions. However, quite a few parents and students knew their scores. "Very little" had been done in regard to remediation for those students who had not mastered the MEAP "essential" objectives.

Parents and students were aware of MEAP test results. Thus, small, rural school districts were doing "quite a bit" of publicity regarding MEAP test results while using these results "some" in the assessment of curricula.

Research Question 1. Are there significant differences among districts with respect to the use of MEAP test results? Eleven items were found to vary significantly in small, rural school districts in the Upper Peninsula. Districts varied in the extent to which the 1989 MEAP results were used to "determine instructional priorities," to change "course content," to alter the "frequency of curriculum revision," and to "identify staff-development needs." Districts varied in the number of inservice activities that were provided to incorporate MEAP objectives and in the sharing of teaching strategies used to increase test scores.

Some districts made greater use of MEAP test scores to "determine the placement of students in 'remedial' programs." Schools differed in terms of teachers' familiarity with MEAP test questions.

The person or persons who interpreted MEAP test results and to whom they were interpreted varied significantly in this study. Some

principals did more interpreting than others. Some school board members and teachers received these reports at their respective meetings; others did not.

Research Question 2. What practices, if any, are intentionally used by small, rural schools in response to the MEAP? In the interviews, several practices surfaced: the administration of the MEAP test by the principal, test preparation, a "pre-MEAP" test, matching of instruction and question format, providing instruction in the weakest areas, a remedial class, and emphasizing to teachers the importance of the MEAP test.

In one district the MEAP was administered by the high school principal in an effort to alert students to the seriousness of the test, thus motivating students to do well. This same district used tests as a criterion for student placement in high- and moderate-ability groups.

Another district provided instruction on test taking at the high school level. At the elementary level, an exercise was done to show students what guessing did to test results.

In Koskitown, a test was administered in the fourth, fifth, and sixth grades. This test contained three questions for each MEAP math objective and a few other items. Students had to pass this test before leaving sixth grade; thus, remediation did take place in the elementary grades.

In at least one district, teachers studied the question format and intentionally present material using this format. A matching of

test vocabulary to instructional vocabulary occurred in this district also.

In most of the districts in which interviews were conducted, MEAP test results were reviewed, looking for weaknesses in the curriculum. If such weaknesses were found, instruction in that subject area was provided.

Several years ago, Koskitown provided a remedial class for those high school students who had not mastered the "essential" objectives in math. Students were able to drop this class when they passed the MEAP test in math.

Teachers in Koskitown believed test results were important. Some thought they were evaluated on the results of these tests. Staff members in this school district were aware of the test content and format; they shared strategies at teacher meetings. Teachers did not want to be "scoring low."

Summary of four small, rural schools. The four small, rural schools in which interviews were conducted were similar in some respects and different in others. All were considered isolated; some were farther from neighboring districts or larger towns than others. All were described as depressed; some communities had more small industries or tourist development than others. All were described as achieving schools; all scored an average of 75 or better on the MEAP. All the schools in which interviews were conducted had concerned and dedicated staff members; educators went about instructing neighborhood children as if they were their own. All educators expressed a desire to improve, through inservice

programs, conferences, or courses; however, accessibility was a problem.

These districts varied in size, organization, and response to test results. Districts varied in student enrollment from 100 to 500. The facilities varied from one to two buildings. The administrative structure could include a superintendent, a high school principal, and an elementary principal; or it might be solely the superintendent. Teachers' meetings could be regular, or they might not be held. Teachers in the district might view themselves as family, or they might be fragmented. Students might be in classes of six or in groups of 30.

District effectiveness tended to be viewed in terms of the success of their graduates. At least one of the districts included test scores in conjunction with student post-high-school success. In this district, educators discussed test scores, looked at test questions, shared strategies, and remediated. However, in other districts, personnel tended to look at the results, recognize that they scored high, and file the test results. Some curriculum revision was done in all districts, but some were involved in revision more than others.

What is clear is that MEAP test results were communicated to parents and students in all districts. All had been successful in scoring above the state average, and this aspect was shared. The question of how these districts would improve beyond the current level was a question that interviewees in some districts raised. An

answer to this question was actively being sought in at least one small, rural school district in the Upper Peninsula of Michigan.

Conclusions

A majority of the small, rural school districts in the Upper Peninsula of Michigan were going about the business of education as they had done in the past--namely, by a very loosely coordinated tinkering with the curriculum by classroom teachers. These modifications, however slight, reinforced the notion that rural schools were providing at least an equal, if not a superior, education to their students.

Clearly, the data indicated that there was a lack of agreement within and between districts. Respondents within districts typically did not agree in their reporting of district practices with regard to the following: standardized tests providing information in assessing effectiveness, the use of MEAP results with other tests, and the use of tests for items not measured by MEAP. Additionally, respondents within districts did not uniformly agree on the use or impact of MEAP test results. Between-district variability was present with regard to the following: the presence of curriculum guides, the use of MEAP to alter instructional content and strategies, the test battery, and the communication of MEAP test results.

What a majority of small, rural school districts lacked was a systematic, technically valid method of developing and reviewing curriculum--in this case, the curriculum-test match. At least half

of the respondents indicated that curricula were not present. In the remaining half, few had examined the matching of the local curriculum objectives to test objectives. The selection of textbooks and classroom materials based on the curriculum was the exception rather than the rule. The tendency to select the textbook and then use it as the curriculum was predominant. While it may appear that the curriculum-test match model was favored by the researcher, it should be acknowledged that the empirical research reviewed did not support any one "best" model. An understanding of the use of criterion-referenced tests was generally lacking.

Within the realm of testing, most teachers did not use standardized test information except to confirm their perceptions about students. They tended to be opposed to teaching to the test objectives. Most wanted to know the "true" score. The distinction between criterion-referenced and norm-referenced tests was not frequently made by these rural educators. There was a tendency to defer testing questions to the counselor, who was designated as the test interpreter and translator of test information. In some cases, the counselor met with teachers and suggested changes that should be made; however, the counselor did not have the authority to enforce such changes.

In summary, tests were given; when the results were received, they were reviewed and then filed. Students in these rural schools scored well, so little was done to alter strategies or content.

However, all of the superintendents who were interviewed indicated a change in the types of new residents in their

communities. Single-parent families and families receiving public assistance were sending new students to school. These families differed markedly from first- or second-generation immigrant families, who valued education and viewed it as the means to success. Teachers indicated that children were "raising themselves," as compared to parents caring for, supporting, and guiding students. According to these teachers, parents were interested in "showcasing" their children to enhance their self-image, rather than being genuinely interested in their youngsters. All of the educators who were interviewed tended to believe that the pool of academic talent was declining. In summary, the clientele in small, rural schools has changed.

Most schools appeared to be doing little to alter their strategies to meet the differing needs of the "new student." When efforts were made, they tended to be by an individual teacher and not coordinated systemwide. They tended to lack the structure and sophistication that would permit an evaluation; the changes seemed unrelated to the test. Most schools just did not subscribe to an explicit, overt curriculum model of teaching specific objectives and testing these objectives. There was a tendency to avoid "teaching as a science" and to cling to "teaching as an art."

In all fairness to these educators in small, rural schools, it was not that they did not want to be the best that they could be, it was just difficult given their degree of isolation from other districts and from the experts. These teachers were as good as any

others; however, the demands of their job differed from those of their urban and suburban counterparts. They indicated a desire to know more, to do things differently, to learn--but at a reasonable sacrifice. They all wanted to be part of an "improving" school, but they believed this could be accomplished only when someone visited and knew their school. These teachers wanted their students to be winners. Thus, the following recommendations are made in the areas of further study and process guidelines. They are as follows:

Recommendations

Based on the findings of this study, the following recommendations for further study are made:

Further Study

1. One district, District 10, along with Koskitown, was found to vary significantly from other districts but was not included in the interview process. Districts like District 10 should be examined to determine what practices are being used to improve students' academic performance. Such information would be helpful as rural districts engage in the school-improvement process.

2. A human resource element appears to be operating when significant variability was present. In both Makivil and Vilmi, administrators did not appear to accept leadership in the curricular domain. This seems appropriate for further study. For example, is this generally true of districts having a student population of less than 300?

3. The small, rural districts appeared to communicate least about tests and curricular issues. Additional studies are needed to confirm and discover the nature and substance of communication in small, rural districts with student enrollments under 300.

4. Additional research should be conducted in small, rural districts to substantiate and investigate whether the variability among respondents when they report district practices is a persistent characteristic.

5. This study should be replicated to include a larger number of districts within and outside Michigan so that findings can be generalized beyond the confines of the Upper Peninsula of Michigan.

Process Guidelines

1. The traditional method of delivering inservice programs does not accommodate the small, rural school well. An on-site visit, a district-tailored workshop, and so on, would be more beneficial to these districts. Television workshops would be an immediate improvement over the current delivery method.

2. Respondents from most districts were satisfied with the number of conferences provided by their local intermediate school district. Their relative proximity afforded greater attendance. However, many voiced a desire to share across grade levels and subject matter rather than bring experts in to share information. Also, a curriculum expert in each intermediate school district would be helpful for districts wishing to bridge the teaching-testing gap.

3. Administrators need to reassess their role as instructional leaders. This role must increase in importance if rural schools are systematically to alter the linkage of curriculum variables. Several avenues to be considered would be conferences, course work, and mentorships.

4. Small, rural school districts must intensify their collaborative efforts so that all will improve through the sharing of effective strategies. They must overcome the current barriers of differing calendars, varying school hours, and local district teacher contracts. Use of technological innovations such as interactive television and electronic blackboards can become a reality as these barriers are eroded.

5. Small, rural districts need to examine their methods of communication in an effort to validate, endorse, and communicate district practices.

APPENDICES

APPENDIX A

SURVEY

RURAL, SMALL SCHOOLS IN THE UPPER PENINSULA



Study conducted by
Faye N. DeMarte
HC Box 16
Drummond Island, Mich.
493-5633

THE USE OF MEAP RESULTS IN RURAL SCHOOL DISTRICTS

Circle a response for each statement which most accurately reflects your perceptions of what occurs within your district.

1. Standardized tests (SAT, ACT, Stanford, Metropolitan, MEAP, etc.) provide important information in assessing your district's effectiveness.

- 1 STRONGLY AGREE
2 AGREE
3 NEUTRAL
4 DISAGREE
5 STRONGLY DISAGREE

2. The following most accurately describes our district:

- 1 MEAP RESULTS ARE USED IN CONJUNCTION WITH OTHER
STANDARDIZED TESTS
2 MEAP RESULTS ARE THE ONLY OUTCOME MEASURES USED IN OUR
DISTRICT
3 MEAP RESULTS ARE NOT USED

3. Our district selects or develops tests for items included in the local curriculum but not measured by MEAP

- 1 YES
2 NO

4. Curriculum guides are available in reading and math.

- 1 YES
2 NO

IF YES, RESPOND USING THE FOLLOWING SCALE:

Not at all		Very Little		Some		Quite A Bit		Extensively	
0	1	2	3	4	5	6	7	8	

Example: A "4" response shows your perception to be "some" but more toward "Quite A Bit".
A "3" response shows your perception to be "some" but more toward "Very Little".

N V S Q E
0 1 2 3 4 5 6 7 8

MEAP OBJECTIVES ARE INCORPORATED INTO OUR
DISTRICT CURRICULUM GUIDES

0 1 2 3 4 5 6 7 8

TEXTBOOKS ARE CHOSEN ON THE BASIS OF OUR
CURRICULUM GUIDES

0 1 2 3 4 5 6 7 8

CURRICULUM GUIDES INFLUENCE CLASSROOM INSTRUCTION

5. Using the eight point scale of "Not at all", "Very Little", "Some", "Quite A Bit", and "Extensively" rate the extent to which you have used the 1989 MEAP results for the following purposes:

	Not at All 0	Very Little 1 2	Some 3 4	Quite A Bit 5 6	Extensively 7 8
A. TO DETERMINE THE GENERAL ACHIEVEMENT LEVEL OF TESTED STUDENTS				N V 0 1 2	S Q E 3 4 5 6 7 8
B. TO INFORM THE SCHOOL COMMUNITY OF THE GENERAL ACHIEVEMENT LEVEL OF TESTED STUDENTS				0 1 2 3 4 5 6 7 8	
C. TO DETERMINE STRENGTHS AND WEAKNESSES IN THE AREA OF MATH				0 1 2 3 4 5 6 7 8	
D. TO DETERMINE STRENGTHS AND WEAKNESSES IN THE AREA OF READING				0 1 2 3 4 5 6 7 8	
E. TO DETERMINE INSTRUCTIONAL PRIORITIES				0 1 2 3 4 5 6 7 8	
F. TO DETERMINE PLACEMENT OF STUDENTS IN "REMEDIAL" PROGRAMS				0 1 2 3 4 5 6 7 8	
G. TO DETERMINE THE NEED FOR NEW PROGRAMS				0 1 2 3 4 5 6 7 8	
H. TO ANALYZE TEACHER PERFORMANCE				0 1 2 3 4 5 6 7 8	
I. TO IDENTIFY STAFF DEVELOPMENT NEEDS				0 1 2 3 4 5 6 7 8	
J. TO PREDICT STUDENTS' FUTURE ACADEMIC SUCCESS				0 1 2 3 4 5 6 7 8	

6. Using the same scale, circle the number that best describes the communication of MEAP test results in your district:

A. THE PRINCIPAL PROVIDES LEADERSHIP IN THE INTERPRETATION OF MEAP TEST RESULTS	N V 0 1 2	S Q E 3 4 5 6 7 8
B. MEAP HAS BEEN DISCUSSED AT TEACHER MEETINGS	0 1 2 3 4 5 6 7 8	
C. MEAP TEST RESULTS HAVE BEEN DISCUSSED AT SCHOOL BOARD MEETING	0 1 2 3 4 5 6 7 8	
D. INSERVICE ACTIVITIES HAVE BEEN PROVIDED TO HELP INCORPORATE MEAP OBJECTIVES INTO INSTRUCTIONAL PROGRAMS	0 1 2 3 4 5 6 7 8	
E. OUR STAFF SHARES INFORMATION REGARDING TEACHING STRATEGIES USED TO INCREASE STUDENT TEST SCORES ON MEAP	0 1 2 3 4 5 6 7 8	

7. Indicate by circling the appropriate number the changes that have been made in your district based on MEAP test results:

		Not At All	Very Little		Some		Quite A Bit		Extensively	
		0	1	2	3	4	5	6	7	8
N	V	S		Q		E				
0	1 2	3 4	5 6	7	8	COURSE OFFERINGS				
0	1 2	3 4	5 6	7	8	COURSE CONTENT				
0	1 2	3 4	5 6	7	8	TEACHING METHODS				
0	1 2	3 4	5 6	7	8	INSTRUCTIONAL METHODS				
0	1 2	3 4	5 6	7	8	PREPARATION FOR TESTS				
0	1 2	3 4	5 6	7	8	BATTERY OF TESTS				
0	1 2	3 4	5 6	7	8	FREQUENCY OF CURRICULUM REVISION				

8. Circle the number that reflects your perception using the previous scale:

A. TEACHERS ARE FAMILIAR WITH MEAP TEST QUESTIONS	N	V	S		Q		E	
	0	1 2	3 4	5 6	7	8		
B. STUDENTS KNOW THEIR MEAP SCORES IN OUR DISTRICT	0	1 2	3 4	5 6	7	8		
C. PARENTS KNOW THEIR STUDENT'S MEAP SCORE	0	1 2	3 4	5 6	7	8		
D. REMEDIAL CLASSES ARE DESIGNED TO TEACH STUDENTS THE MASTERY OF SPECIFIC MEAP OBJECTIVES	0	1 2	3 4	5 6	7	8		

Please indicate your position:

- 1 ADMINISTRATOR
- 2 ELEMENTARY TEACHER
- 3 JUNIOR HIGH TEACHER
- 4 HIGH SCHOOL TEACHER

() I WOULD LIKE A COPY OF THIS STUDY'S FINDINGS

ADDITIONAL COMMENTS _____

SAY
"YES"
TO RESEARCH
OF
THE U.P.
FOR
THE U.P.

APPENDIX B

LETTER OF PERMISSION TO CONDUCT STUDY

MICHIGAN STATE UNIVERSITY

COLLEGE OF SOCIAL SCIENCE • OFFICE OF THE DEAN
205 BERKEEY HALL

EAST LANSING • MICHIGAN • 48824-1111

April 17, 1990

IRB# 90-158

Faye N. DeMarte
HC 52 Box 16
Drummond Island, MI 49726

Dear Ms. DeMarte:

RE: "THE USE OF THE MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM
IN SMALL, RURAL SCHOOL DISTRICTS IN THE UPPER PENINSULA OF
MICHIGAN IRB# 90-158"

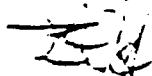
The above project is exempt from full UCRIHS review. I have reviewed the proposed research protocol and find that the rights and welfare of human subjects appear to be protected. You have approval to conduct the research.

You are reminded that UCRIHS approval is valid for one calendar year. If you plan to continue this project beyond one year, please make provisions for obtaining appropriate UCRIHS approval one month prior to April 17, 1991.

Any changes in procedures involving human subjects must be reviewed by UCRIHS prior to initiation of the change. UCRIHS must also be notified promptly of any problems (unexpected side effects, complaints, etc.) involving human subjects during the course of the work.

Thank you for bringing this project to our attention. If we can be of any future help, please do not hesitate to let us know.

Sincerely,



John K. Hudzik, Ph.D.
Chair, UCRIHS

JKH/sar

cc: S. Moore

APPENDIX C

COVER LETTER

April 18, 1990

Dear _____,

Your school has been selected randomly to participate in a voluntary study designed to find out if and how small, rural schools in the Upper Peninsula use Michigan Educational Assessment Program (MEAP) results.

In developing the MEAP, the Michigan Department of Education intended for local school districts to use the assessment results to evaluate their curriculum. The data collected by this survey should answer the question of whether MEAP test results are used, how they are used, and what practices small, rural districts use to incorporate MEAP into their districts.

Enclosed is a survey and envelope. This survey can be completed in fifteen minutes. Please place the completed survey in the envelope and return it to [the superintendent's secretary]. I will visit your district during the week of April 23 to collect the surveys.

All answers will be kept confidential. Participation in this study is completely voluntary. You indicate your voluntary agreement to participate by completing and returning this survey. On request, a copy of the final report will be sent to your district upon completion of this study.

If you have any questions, please call me at (906) 493-5633. Thank you in advance for your assistance.

Sincerely,

Faye N. DeMarte

Enclosure

APPENDIX D

TABLES

Table D.1.--Percentages: Curriculum guides are available in reading and math.

District	Presence of Curriculum Guides		
	Yes	Of those indicating the presence of guides, then -	
1	53.8		
2	62.5	Objectives in guides	3.92
3	64.3		
4	63.6		
5	0.0	Textbooks chosen as	
6	33.3	basis of guides	4.10
7	50.0		
8	54.5	Guides influence	
9	25.0	instruction	4.36
10	80.0		

Table D.2.--Means for all variables on survey instrument.

Variable		Mean
Q1	Test's importance	2.247
Q2	MEAP usage with other tests	1.204
Q3	Test development	1.527
Q4	Curriculum guides' presence	1.478
Q4A	MEAP objectives in guides	3.918
Q4B	Textbook and guides linkage	4.102
Q4C	Guides influence instruction	4.632
Q5A	Determine general achievement level	3.967
Q5B	Inform community	3.870
Q5C	Determine strengths/weaknesses in math	4.142
Q5D	Determine strengths/weaknesses in reading	4.163
Q5E	Instructional priorities	3.870
Q5F	Placement in remedial programs	2.815
Q5G	Need for new programs	2.760
Q5H	Analyze teacher performance	1.521
Q5I	Identify staff needs	2.591
Q5J	Predict future success	2.397
Q6A	Principal provides leadership	2.957
Q6B	Discussed at teachers' meetings	4.361
Q6C	Discussed at board meetings	4.478
Q6D	Inservice activities	2.681
Q6E	Staff shared information	2.872
Q7A	Course offerings	2.478
Q7B	Course content	3.413
Q7C	Teaching methods	3.365
Q7D	Instructional methods	3.365
Q7E	Preparation for tests	3.282
Q7F	Battery of tests	2.428
Q7G	Frequency of curriculum revision	2.666
Q8A	Teachers familiar with questions	4.127
Q8B	Students know their score	5.527
Q8C	Parents know student's score	5.693
Q8D	Influence design of classes	2.233

Table D.3.--Means for significant contrasts.

Question	Schools	Mean
5e	Unknown	
5f	2, 7, 8, 9, & 10	1.8032
	All the rest	3.6587
5i	10	4.8000
	1, 5, & 9	1.6230
6a	5 & 9	0.4643
	All the rest	3.4486
6b	5 & 9	2.4732
	All the rest except 1	4.8676
6c	Unknown	
6d	1, 5, & 7	1.0843
	All the rest	3.3797
	10	6.5000
	All the rest	2.2678
6e	5, 8, & 9	3.3468
	All the rest	1.3701
	10	5.6000
	All the rest	2.4375
7b	5	1.0000
	All the rest	3.6640
7g	4 & 10	3.9045
	1 & 3	1.1952
	4 & 10	3.9045
	1, 3, & 5	1.5879
8a	1, 5, & 9	3.0485
	4 & 10	5.4227

Table D.4.--Statistically significant contrasts within districts
for questions that are statistically significant between
districts.

Question	School	Number of Respondents
5e	6	6
5f	Makivil	7
6a	Koskitown	12
	8	11
6b	8	11
6d	Makivil	7
6e	3	14
7g	Vilmi	8
8a	3	14
	6	6

BIBLIOGRAPHY

BIBLIOGRAPHY

- Airasian, P. W. (1979). Educational measurement and technological bases underlying minimum competency testing. In P. W. Airasian, G. F. Madaus, & J. J. Pedula (Eds.), Minimum competency testing (pp. 33-48). Englewood Cliffs, NJ: Educational Technology Publications.
- Airasian, P. W., Madaus, G. F., & Pedula, J. J. (Eds.). (1979). Minimum competency testing. Englewood Cliffs, NJ: Educational Technology Publications.
- Aquino, A. F. (1975). Attitudes of fourth and seventh grade teachers toward the 1974-75 Michigan Educational Assessment Program and their use of the results. Ph.D. dissertation, Wayne State University.
- Barker, B. O. (1985). Research focus on rural schools in Oklahoma. El Reno, OK: Annual Conference of the Communities United for Rural Education. (ERIC Document Reproduction Service No. ED 256 536)
- Barker, B., & Muse, I. (1983). Research findings on K-12 and 1-12 rural school districts in the United States. Manhattan, KA: Annual Conference of the Rural Education Association. (ERIC Document Reproduction Service No. ED 234 973)
- Barker, B. W., Muse, I., & Smith, R. B. (1983). A report on rural education in Arkansas. (ERIC Document Reproduction Service No. ED 231 583)
- Blau, T. H. (1979). Minimum competency testing, psychological implications for students. In P. W. Airasian, G. F. Madaus, & J. J. Pedula (Eds.), Minimum competency testing. Englewood Cliffs, NJ: Educational Technology Publications.
- Bleecher, H. (1975). The authoritativeness of Michigan's accountability program. Ph.D. dissertation, Michigan State University.
- Brezekinski, E. J. (1976). Testing in Michigan: A twenty-year perspective. Ann Arbor: University of Michigan, Bureau of School Services.
- Brown v. Board of Education, 347 U.S. 483 (1954).

- Bushaw, W. J. (1988). Assessment testing: Implementation and utilization. Ph.D. dissertation, The University of Michigan.
- Buttram, J. L., & Carlson, R. V. (1983). Effective school research: Will it play in the country? Proceedings of the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada. (ERIC Document Reproduction Service No. ED 235 561)
- Cohen, S. A. (1987). Instructional alignment: Searching for a magic bullet. Educational Researcher, 16, 16-19.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A., Weinfeld, F. D., & York, R. (1966). Equality of educational opportunity. Washington, DC: U.S. Government Printing Office.
- Cooley, W. W., & Leinhard, T. G. (1980). The instructional dimensions study. Educational Evaluation and Policy Analysis, 2, 7-25.
- Darling-Hammond, L., & Wise, A. (1985). Beyond standardization: State standards and school improvement. Elementary School Journal, 85(3), 315-336.
- Debra P. v. Turlington, 730 S. 2nd 1405, 11th Cir. (1984).
- DeYoung, A. (1987). The status of American rural education research: An integrated review and commentary. Review of Educational Research, 57(2), 123-148.
- Dillman, D. A. (1978). Mail and telephone surveys: The total design method. New York: John Wiley & Sons.
- Donovan, D. (1982). All children can learn--Schools make a difference. Michigan School Board Journal, 29(4), 8-11.
- Edelman, J. (1981). The impact of the mandated testing program on the classroom practices: Teachers' perspectives. Education, 102(1), 56-59.
- English, F. (1978). Quality control in curriculum development. Arlington, VA: American Association of School Administrators.
- English, F. (1987). The principal as master architect of curricular unity. National Association of Secondary School Principals Bulletin, 72, 35-42.
- Evans, M. C. (1985). Is state legislation a catalyst for change? Phi Delta Kappan, 66(7), 501-503.

- Farfaring, J. E. (1988). A study of the correlation between key traits of effective school research and the Michigan Educational Assessment Program. Dissertation Abstracts International, 26, 362. (University Microfilms No. ADG13-33085, 72)
- Fidler, R. A. (1988). Factors explaining success and lack of success on the Michigan Educational Assessment Program science assessment. Dissertation Abstracts International, 49, 3601. (University Microfilms No. 89-0697, 130)
- Fisher, T. H., & Roth, R. W. (1972). A descriptive study of local districts' reactions to the Michigan Educational Assessment Program. Michigan Association of School Boards Journal, 12, 13-15.
- Floden, R. E., Porter, A., Alford, L., Freeman, D., Irwin, S., Schmidt, W., & Schwille, J. (1988). Instructional leadership at the district level. Educational Administration Quarterly, 24(2), 96-124.
- Floden, R. E., Porter, A., Schmidt, W., Freeman, D., & Schwille, J. (1981). Responses to curriculum pressures: A policy-capturing study of teacher decisions about content. Journal of Educational Psychology, 73(2), 129-141.
- Gjelten, T. (1982). A typology of rural school settings. Washington, DC: U.S. Department of Education, Rural Education Seminar. (ERIC Document Reproduction Service No. ED 215 858)
- Gould, J. C. (1980). Why the AFT support standardized testing. New Directions for Testing and Measurement, 16, 53-57.
- Hall, G., Jaeger, R. M., Kearney, C. P., & Wiley, D. E. (1987). Alternatives for a national data system on elementary and secondary education. International Journal of Educational Research, 11(4), 453-502.
- Haney, W. (1984). Testing reasoning and reasoning about testing. Review of Educational Research, 54(4), 597-654.
- Haney, W. (1985). Making testing more educational. Educational Leadership, 43(2), 4-13.
- Haney, W., & Madaus, G. F. (1979). Making sense of the competency movement. In P. W. Airasian, G. F. Madaus, & J. J. Pedula (Eds.), Minimum competency testing. Englewood Cliffs, NJ: Educational Technology Publications.

- Hornbeck, M. (1989, June 11). The MEAP and per-pupil spending. Detroit Free Press, p. 14a.
- House, E. R. (Ed.). (1973). School evaluation: The politics and process. Berkeley: McCutchan.
- House, E., Rivers, W., & Stufflebeam, D. (1974). An assessment of the Michigan accountability system. Lansing: Michigan Education Association.
- Idol, L., Nevin, A., & Paolucci-Whitcomb, P. (1986). Models of curriculum-based assessment. Rockville, MD: Aspen Publications.
- Jaeger, R. M., & Tittle, C. (Eds.). (1980). Minimum competency achievement testing: Motives, models, measures, and consequences. Washington, DC: American Educational Research Association.
- Jencka, C. G. (1990). A longitudinal investigation of school administrators' practices and attitudes related to the use of the Michigan Educational Assessment Program test results. Ph.D. dissertation, Michigan State University.
- Jess, J. D. (1984). Rural education 1984: Issues and impacting forces. A local perspective. (ERIC Document Reproduction Service No. ED 251 253)
- Kelleghan, T., Madaus, G. F., & Airasian, P. W. (1982). The effects of standardized testing. Boston: Kluwer-Nijoff Publishing.
- Kirst, M. (1987). Curricular leadership at the state level: What is the new focus? National Association of Secondary School Principals Bulletin, 72(505), 8-14.
- Klein, K. (1984). Minimum competency testing: Shaping and reflecting curricula. Phi Delta Kappan, 65, 565-567.
- Kohr, R. L. (1983). An examination of community type differences in the educational quality assessment data. (ERIC Document Reproduction Service No. ED 244 972)
- LaMahiew, P., & Leinhardt, G. (1985). Overlap: Influencing what's taught. A process model of teachers' content selection. Journal of Classroom Interaction, 21(1), 2-11.
- Lazarus, M. (1981). Goodbye to excellence: A critical look at minimum competency testing. Boulder, CO: Westview Press.

- Lazarus, M. (1982). Evaluating educational programs. Arlington, VA: American Association of School Administrators.
- Lessinger, L. M. (1970). Every kid a winner: Accountability in education. New York: Simon & Schuster.
- Lessinger, L. M., & Tyler, R. W. (Eds.). (1971). Accountability in education. Worthington, OH: Charles A. Jones.
- Loustaunau, M. (1975). Small rural schools can have adequate curriculums. University Park: New Mexico State University. (ERIC Document Reproduction Service No. ED 100 559)
- McClurkin, W. D. (1970). Rural education in the United States. University Park: New Mexico State University. (ERIC Document Reproduction Service No. ED 043 408)
- McCormick, R. (1978). The controversial development of the Michigan Educational Assessment Program. Ph.D. dissertation, Michigan State University.
- Madaus, G., Kelleghan, T., Rakaw, E. A., & King, D. J. (1979). The sensitivity of measures of school effectiveness. Harvard Education Review, 49, 207-230.
- Marshall, D. G. (1986). From rhetoric to action: Approaches to small schools for the coming decade. Proceedings of the Annual Meeting of the Canadian National Symposium on Small Schools, Atkoka, Ontario, Canada. (ERIC Document Reproduction Service No. ED 276 106)
- Mathison, S. M. (1987). The perceived effects of standardized testing on teaching and curricula. Dissertation Abstracts International, 49, 35. (University Microfilms No. ADG88-03130, 252)
- Mehrens, W. A. (1979). The technology of competency measurement. In R. Ingle, M. Carroll, & W. Gephart (Eds.), Assessment of student competence. Bloomington, IN: Phi Delta Kappan.
- Mehrens, W., & Phillips, S. E. (1986). Detecting impacts of curricular differences in achievement data. Journal of Educational Measurement, 23(3), 185-196.
- Michigan Association of Professors of Educational Administration. (1972). The Michigan Educational Assessment Program: An abstract. East Lansing: Author.
- Michigan Department of Education. (1972). Analyzing Michigan's educational progress. Lansing: Author.

- Michigan Department of Education. (1985). Pencils down. A guide for using and reporting test results. Lansing: Author.
- Mika, W. J., Jr. (1982). An NEA-VEA response. New Directions for Testing and Measurement, 16(53), 59-64.
- Muse, I. D., Barker, B. O., & Smith, R. B. (1983). Rural education in Colorado: Recent research findings. Proceedings of the Annual Colorado Rural and Small Schools Conference, Fort Collins, CO. (ERIC Document Reproduction Service No. ED 231 573)
- Music, P. H. W. (1982). An analysis of the relationship between selected instructional materials and the Michigan Educational Assessment Program objectives for the primary grades. Dissertation Abstracts International, 43, 2886. (University Microfilms No. ADG83-03831, 144)
- Nachtigal, P. M. (1980). Improving rural schools. Denver, CO: Education Commission of the States. (ERIC Document Reproduction Service No. ED 192 992)
- National Education Association. (1979). Handbook.
- National School Boards Association. (1977). Standardized testing.
- Orlich, D., & Ratcliff, J. L. (1977). Coping with the myth of accountability. Educational Leadership, 34(4), 246-251.
- Pedula, J. J., & Reidy, E. F. (1979). The rise of the minimal competency testing movement. In P. W. Airasian, G. F. Madaus, & J. J. Pedula (Eds.), Minimal competency testing (pp. 23-33). Englewood Cliffs, NJ: Educational Technology Publications.
- Popham, W. J. (1977). Customized criterion-referenced tests. Educational Leadership, 34(4), 258-259.
- Popham, W. J. (1985). Measurement-driven instruction It's on the road. Phi Delta Kappan, 66(9), 628-634.
- Porter, A. C. (1983). The role of testing in effective schools. American Education, 19(1), 25-28.
- Porter, A. C. (1988). Objective indicators: Objective data or political tool. Phi Delta Kappan, 69(8), 503-505.
- Porter, A. (1989). A curriculum out of balance: The case of elementary school mathematics. Educational Researcher, 18(5), 9-15.

- Porter, J. W. (1976). The virtues of a state assessment program. Phi Delta Kappan, 58, 665-668.
- Public Sector Consultants. (1989). Public sector report.
- Resnick, D., & Resnick, L. (1985). Standards, curriculum, and performance: Historical and comparative perspective. Educational Researcher, 15, 5-20.
- Rudman, H. (1977). The Michigan Educational Assessment Program, 1976-77: Its meaning to school boards. Michigan School Board Journal, 24(7), 10-11.
- Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? Phi Delta Kappan, 62(9), 631-634.
- San Antonio Independent School District v. Rodrigues. 96 Cal. Rptr. 601 (1971).
- Schalock, D., Felding, G., Schalock, M., Erickson, J., & Brott, M. (1985). Integrating teaching and testing with program management. Educational Leadership, 43(2), 55-58.
- Sher, J. P. (1977). Public education in sparsely populated areas of the United States. Washington, DC: Department of Health, Education, & Welfare, National Institute of Education.
- Sher, J. P. (1978). Revitalizing rural education: A legislator's handbook. Washington, DC: Department of Health, Education, & Welfare, National Institute of Education. (ERIC Document Reproduction Service No. ED 168 750)
- Shinsky, E. J. (1983). The appropriateness of the criteria for inclusion of special education students in the Michigan Educational Assessment Program at the fourth grade level. Dissertation Abstracts International, 45, 814. (University Microfilms No. ADG84-07238, 203)
- Simon, R. F. (1982). Teachers' perceptions of students' performance on Michigan Educational Assessment Program mathematics tests. Dissertation Abstracts International, 43, 2324. (University Microfilms No. AADG82-27818, 155)
- Smedley, S. L. (1987). Standardized achievement testing in education and the 1987 Washington State Testing Program. Ed.D. dissertation, Brigham Young University.
- Statistical Analysis Systems. (1985). North Carolina: SAS Institute.

- Steele, D. J., Jr. (1976). An investigation of administrative practices and attitudes related to the use of Michigan Educational Assessment Program test results. Ph.D. dissertation, The Ohio State University.
- Stiggins, R. J. (1985). Improving assessment where it means the most: In the classroom. Educational Leadership, 43(2), 69-74.
- Thelander, J. E. (1979). An examination of the relationship of the Michigan Educational Assessment Program to local school district curriculum decision making. Ph.D. dissertation, The University of Michigan.
- Tyler, R. W. (1971). Accountability in perspective. In L. M. Lessinger (Ed.), Accountability in education (pp. 1-80). Worthington, OH: Charles A. Jones Publishing House.
- Walstad, W. B. (1984). Analyzing minimum competency test performance. Journal of Educational Research, 77(5), 261-266.
- Watson, R., & Schmalgemeir, W. L. (1970). Michigan schools: The organization and management of their testing programs. Ann Arbor: The University of Michigan, Bureau of School Services.
- Wise, A. (1979). Legislated learning: The bureaucratization of the American classroom. Berkeley: University of California Press.
- Womer, F. B. (1959). Testing programs in Michigan schools. Ann Arbor: The University of Michigan, Michigan School Testing Services.
- Womer, F. D. (1981). State level testing: Where we have been may not tell us where we are going. In D. Carlson (Ed.), New directions for testing and measurement: Vol. 10. Testing in the states: Beyond accountability (pp. 1-12). San Francisco: Jossey-Bass.