DOCUMENT RESUME

ED 104 963 TM 004 549

AUTHOR	Abbott, Muriel M.; And Others
TITLE	The Development and Management of a Pilot Project for a Reading Assessment System.
PUB DATE	Apr 73
NOTE	56p.; Symposium papers presented at the Annual Meeting of the American Educational Research Association (New Orleans, April 1973)
EDRS PRICE	MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS	*Behavioral Objectives; Criterion Referenced Tests; Decision Making; Educational Objectives; Evaluation; Feedback; *Formative Evaluation; *Individualized Instruction; Individualized Programs; Individualized Reading; *Integrated Curriculum; Models; Norm Referenced Tests; Reading Achievement; Reading Instruction; *Reading Programs; Summative Evaluation; Test Construction; Testing
IDENTIFIERS	Harcourt Brace Jovanovich; *Skills Monitoring System

ABSTRACT

The objectives of the symposium were to provide an example of the institution of a systems approach to the solving of an educational problem and to relate this approach to current measurement issues such as criterion-referenced interpretation of test scores and individualization of assessment, diagnosis and instruction. A pilot project in reading assessment was developed and implemented. It consisted of an integrated package, all components being interrelated through behavioral objectives. The system included a set of behavioral objectives, instructional programs designed to lead to their achievement, a means of measuring the extent to which objectives are achieved, and provision for feedback to decision makers. Representatives of the test publisher discussed the design and production of the assessment system. Representatives of the local educational agencies discussed implementation and function of the pilot project in the actual school situations. The educational importance of the symposium lies in the fact that although the issues faced and solutions derived are discussed with reference to a particular pilot project in reading, they are applicable to other instructional situations. (Author/BJG)

THE DEVELOPMENT AND MANAGEMENT OF A PILOT PROJECT FOR A READING ASSESSMENT SYSTEM

Muriel M. Abbott, Organizer
Ronald E. Banks, Chairman
Lenore Ri[illegible]er, Discussant
Muriel M. Abbott
Richard C. Benjamin
Kenneth L. Carlson
Barrie Wellens
G. Richard Zubulake

Symposium presented at the Annual Meeting of the American Educational Research Association in New Orleans, 1973.

Published by:
Harcourt Brace Jovanovich, Inc.
New York, April 1973

THE DEVELOPMENT AND MANAGEMENT OF A PILOT PROJECT FOR A READING ASSESSMENT SYSTEM

Muriel M. Abbott, Organizer
Ronald E. Banks, Chairman
Lenore Ringler, Discussant
Muriel M. Abbott
Richard C. Benjamin
Kenneth L. Carlson
Barrie Wellens
G. Richard Zubulake

Symposium presented at the Annual Meeting of the American Educational Research Association in New Orleans, 1973.

Published by:
Harcourt Brace Jovanovich, Inc.
New York, April 1973

TABLE OF CONTENTS

The Development and Management of a Pilot Project for a Reading Assessment System

General Abstract

The objectives of the symposium were to provide an example of the institution of a systems approach to the solving of an educational problem and to relate this approach to current measurement issues such as criterion-referenced interpretation of test scores and individualization of assessment, diagnosis and instruction. A pilot project in reading assessment was developed and implemented. It consisted of an integrated package, all components being interrelated through behavioral objectives. The system included a set of behavioral objectives, instructional programs designed to lead to their achievement, a means of measuring the extent to which objectives are achieved, and provision for feedback to decision makers. Representatives of the test publisher discussed the design and production of the assessment system. Representatives of the local educational agencies discussed implementation and function of the pilot project in the actual school situations.

The educational importance of the symposium lies in the fact that although the issues faced and solutions derived are discussed with reference to a particular pilot project in reading, they are applicable to other instructional situations. One factor of great importance today is individualization of instruction. The HBJ Skills Monitoring System in Reading is designed to facilitate this in the classroom. The increasing emphasis on the use of behavioral objectives in school situations makes very timely an example of their role in an assessment system in process in three different school situations.

RATIONALE UNDERLYING A TEST PUBLISHER'S APPROACH TO MEASUREMENT IN A SKILLS MONITORING SYSTEM IN READING

Muriel M. Abbott, Ph.D.
Managing Editor, Test Department
Harcourt Brace Jovanovich, Inc.

Increasingly, attention is being paid to the humanizing and individualizing of educational programs and educational measurement. The Skills Monitoring System in Reading that will be discussed today represents an attempt to implement this approach. In the past, much assessment of pupil progress was for the purpose of administrative decision-making. This resulted in insufficient attention to teacher and pupil needs. Recently, however, there has been increasing awareness of the need for measurement of pupil progress that will provide information conducive to teacher decision-making as well as provide a productive learning experience for the individual pupil. This need has brought about a demand for pupil assessment that can supplement the traditional type of assessment program; a program that all too often superimposed assessment onto instruction rather than integrated assessment with instruction. One supplementary type of assessment that is desirable is one that will constitute an integral part of the classroom instructional program and will, in combination with that instructional program constitute a classroom management system.

Demand for change is no stranger to the field of education, and measurement has always reflected educational change. In fact, measurement has often been in the forefront of educational controversy related to change.

Recent controversy over testing concerns not so much what is being measured, as how it is measured, and not so much that measurement takes place, as the use of its results. For example, there is agreement that word knowledge is important; the disagreement centers around what words are to be measured; that is, the relevance of a word in the context of an individual's experiential background. The controversy concerns not so much that an individual's performance be measured, as how the results of that measurement be interpreted and used. The issue of interpretation basically is concerned with whether the measurement should provide answers to the question, "How does an individual compare with others?", as in the case of a normative interpretation, or, whether it should provide answers to the question, "In what way does the individual behave?" or "What does he know?" as in criterion interpretation. The issue of use basically is concerned with summative or formative evaluation; that is, should the results be used to determine status - where the individual is, - or should they be used for prescriptive purposes - what should be done for him next. The issues of interpretation and use are reflected in the objective - or criterion-referenced testing movement as well as in the movement to integrate assessment with the instructional program.

In the past, most educational measurement was for the purpose of summative evaluation, to determine the status of an individual or group. This was especially true of system-wide testing programs. In system-wide testing programs, the commercially published instrument most widely used to measure status was the norm-referenced test. These instruments reported results for an achievement area as a whole, but not for specific

components of that area. As a basis for formulating administrative decisions, this was often sufficient, but as a basis for individual teacher decisions, it left much to be desired. The failure of these instruments to identify specific components of a subject area meant that the usefulness of these instruments for purposes of formulative evaluation was limited to the diagnosis of strength or weakness in an achievement area as a whole. Test results of this type failed to provide a teacher with the diagnostic information needed for the planning of appropriate instructional activities. The same, unfortunately, was often true of teacher-made tests which, although measuring a more limited area of achievement, failed to provide this kind of specific diagnostic information and served primarily as a basis for rank-ordering students. The limitations of norm-referenced tests for formative purposes have been somewhat overcome by greater specificity in the reporting of results made possible through the availability of item-analysis data for specific items measuring defined content within an achievement area.

Recently, commercially-published and independently-developed criterion-referenced instruments have been used for purposes of summative evaluation. Typically, these tests measure and report achievement of many specific objectives within a particular subject area. Insofar as these objectives are more specifically defined than the content areas of the norm-referenced tests and the results reported with correspondingly greater specificity, to that extent do the criterion-referenced instruments have greater potential for purposes of formative evaluation.

It should be noted, however, that such is the case only when both the criterion and norm-referenced instruments are valid measures of the instructional program. There is nothing inherent in either type of instrument that ensures its content validity or relevance for a particular program.

However, both the norm-referenced and criterion-referenced instruments that have been developed and used for summative purposes tend to be of limited usefulness for formative purposes. Because these typically are designed to be given only once during a school year, they provide diagnostic information conducive to prescriptive action at that point in time only. Therefore, they do not provide the kind of continuous diagnostic information needed by a teacher in a classroom situation. The typical commercially published diagnostic instrument also has limitations for formative evaluation. These tests are designed primarily to diagnose the reading-skill deficiencies of poor readers. Their purpose is not to identify specific skill needs of students at all levels of achievement. Furthermore, because they, too, typically are given only once or twice during a school year, they do not provide the kind of continuous diagnostic information required by a teacher in a classroom situation.

Another deficiency of most commercially published assessment instruments, both norm and criterion-referenced, is their failure to provide immediate feedback of results to teacher and student. The two-week or more interval necessary for scoring, processing, and reporting back test results is

acceptable when the results are to be used for summative purposes or to provide a basis for administrative decisions. However, if the results are to be used to provide diagnostic information to aid the teacher in prescribing appropriate instruction, then any such time interval constitutes a considerable obstacle to prompt prescription and its implementation. A very real limitation of almost all tests, is their inability to provide a productive learning experience for the student. In the case of a norm-referenced test, information provided to the student may consist only of where he stands relative to others in some global area of achievement. In the case of an objective- or criterion-referenced test, the information given to the student may be more specific in that it may identify his areas of weakness or strength within the achivement area. But rarely does the test itself constitute a teaching situation. From an educational point of view, there are considerable advantages to a test that can tell a student immediately not only whether he knows something or not, but what the correct information is, if he does not know it.

Another very real problem is testing time. The norm-referenced instrument requires considerable time for the assessment of pupil performance in one content area. Even more time, however, is required for administration of a criterion-referenced test which measures a great many specific objectives in a particular subject area. The typical procedure has been to measure all of a content area or all of the objectives in a subject area in one or two lengthy sessions. Needs were diagnosed based upon the results of that testing. There was no provision for systematic follow-up assessment.

Continuous follow-up diagnosis by means of the comprehensive instrument would be impractical and wasteful. At any one time, for any particular pupil, instruction can be given in only selected areas of identified weakness. Therefore, the repeated assessment of all areas to measure progress in the few would be inefficient. A more useful approach is to administer a comprehensive instrument at the beginning of a relatively long time period, such as a school year, to identify areas of weakness. Then provide for the continuous systematic diagnosis of need by means of short tests measuring performance in the specific areas where instruction was applied.

If a teacher is to provide appropriate instruction to meet specific needs, an instructional support system is desirable. Therefore, a classroom management system should include guidelines to sources of instructional materials that are appropriate for meeting identified needs.

If assessment of pupil progress is to be of maximum benefit both to teacher and student, it must be integrated with the instructional program. For the teacher, the assessment must provide for continuous diagnosis of individual needs with immediate feedback of identified needs in a form conducive to the prescription of appropriate instruction. Because individual needs differ not only in their nature at the time of assessment but also in their rate of change after assessment, the system must be sufficiently flexible to handle these differences. The assessment should be so closely related to the instructional program that as needs are diagnosed, there is readily available a prescription for appropriate

instructional materials to meet diagnosed needs. For the student, the assessment should provide a learning experience. He should discover immediately not only what he knows or does not know but also should receive reinforcement of the correct answer or information as to what the correct answer is.

The HBJ Skills Monitoring System in Reading is designed to provide for the meeting of teacher and student needs through the integration of instructional program and assessment within the classroom situation. It is based upon objective-referenced testing and a test-teach-test method of instruction. In constructing any such system there must be a design and development of behavioral objectives, an instructional program or materials leading to the achievement of each objective, a system of measurement of the achievement of each objective, and provision for feedback of results. The basic component of the system is the objective and all phases of the system are interrelated through the objectives.

The system provides for the diagnosis of student needs by means of a comprehensive Skill Locator measuring all the objectives. It is administered at the beginning of the school year. It provides for continuous diangostic testing by means of short Skill Mini-Tests each measuring a single objective. The results from these tests enable the teacher to diagnose need and to prescribe appropriate instruction to meet changing needs as they are identified. Guidelines are being developed to identify instructional materials appropriate for each Skill Mini-Test. The Skill

Mini-Tests are designed so that the student receives immediate feedback as to whether his response to an item is correct or not, and if not, what the correct answer is. The specifics of the system will be described in the paper by Dr. Wellens.

A PUBLISHER'S IMPLEMENTATION OF CRITERION-REFERENCED MEASUREMENT IN A SKILLS MONITORING SYSTEM IN READING

Barrie Wellens, Ph.D.
Executive Editor, Test Department
Harcourt Brace Jovanovich, Inc.

This paper will deal with the technical aspects confronting a test publisher in the development of a classroom management system in reading. The first step in the creation of the system was to develop an item bank based on behavioral objectives. A set of about 300 objectives in Grades 1 through 6 was developed by measurement and reading experts, and then assigned to grade levels with the aid of classroom teachers. (Many objectives apply to two or three grades.)

The objectives were categorized into three major areas of reading: Word Identification, which involves such topics as vowels, consonants, prefixes, suffixes, word forms, and phonetic parts; Comprehension, which ranges from word meaning to main idea to critical reading skills; and Study Skills, covering the ability to follow directions, dictionary usage, locational and organizational skills, and graphic materials.

Item types were devised to measure each objective at each designated grade level. Objectives which could not be measured appropriately by group-administered paper and pencil items were not forced into this mold. Some objectives, for example, called for an open-ended oral response, some for task performance, and some for audio-visual equipment.

The major guideline in writing item specifications, however, was that paper and pencil, group-administered items were to be used wherever they were as efficient as other item formats in measuring an objective fully. Testing time had to be used effectively, and administration and scoring procedures had to be made as simple as possible.

When presentation-response modes were resolved, item specifications were written and sample items constructed. Whenever an objective was to be measured at more than one grade level, a different set of item specifications was written for each grade level. Where an objective could be measured in different ways, the item specifications contained several sample items to show a variety of techniques. More than 2500 test items were then written by professional item writers and edited by a team of test editors and a qualified reading consultant.

Once the item bank had been created, the classroom management system was designed. The principal component of the system is the Skill Mini-Test. This is a one-page power test designed for the test-teach-test method of instruction. This method involves testing to see which skills must be taught to each individual, providing instruction, and testing again to see whether the instruction was effective. If the pupil achieves mastery, the teacher can then go on to the next area of instruction; if the pupil does not achieve mastery, the teacher should provide more intensive coverage of the skill and then test again.

Matters to be considered in the construction of these tests included

item format, testing time, number of items needed to measure each objective, and the criterion for mastery. It was decided that all tests should be self-administering so that any test could be taken at any time without teacher involvement and without special equipment. In this way, a pupil could be taking one test while his neighbor took another. In order to measure as many objectives as possible, items from the bank were converted to the multiple-device format where this was appropriate.

The questions of testing time and number of items required to test each objective go hand in hand. In order not to sacrifice reliability, which is especially important where individual performance is being assessed, it was decided that six items be used to measure each objective. (This has since been revised to twelve items.) Although designed to be a power test with no time limits, most pupils can complete a Skill Mini-Test in five to fifteen minutes.

Determination of the criterion for mastery still remains a problem. Until data gathered from the pilot testing of this system have been analyzed, the passing score of 75% to 80% of items correct is being used.

In building the Mini-Tests, a method of providing immediate feedback to both pupil and teacher was sought. Of the many self-scoring devices investigated, the 3-M Company's latent image process was selected since it had the most desirable features. With this technique, the examinee marks his answers directly on the test paper with a chemically treated crayon. When he fills in the answer box, a "Yes" or a "No" appears in the box. If he gets a "Yes," he is told to go on to the next question.

If a "No" appears, he is told to stop and think, and try again until he gets a "Yes." This provides him not only with immediate reinforcement, but also with a learning experience since he sees what the correct answer is. The teacher also benefits from this method if she inspects the wrong answers chosen by a pupil and the number of attempts he had to make to get the correct answer.

The Mini-Tests are scored by counting the number of "Yesses" on the first try. (The teacher has a class record sheet which she keeps up to date by scoring each Mini as the pupil brings it to her and recording whether he has passed or failed the Mini.) The teacher should use the scores not only as a guide to instruction, but as an aid in reporting to parents. Results may accompany pupils when they move from one classroom to another, from one school to another, or from grade to grade. This information may also be processed by computer to provide periodic summary data for individuals or groups.

It would be extremely time-consuming to administer each Mini-Test to each pupil in order to determine initially each pupil's needs. Therefore, another major component of the system, the Skill Locator, has been constructed to measure Word Identification, Comprehension, and Study Skills. The Skill Locators are survey tests containing the two items which best predict each skill measured in the Mini-Tests. The teacher organizes her class into instructional groups based on the Skill Locator results. An individual or group is entered into an instructional program in one of the identified areas of weakness. Then, when the

teacher or the pupil believes the pupil has mastered the particular area, the appropriate Mini-Test is given. It should be noted that instructional decisions are based upon performance on clusters of related skills rather than on just the two Skill Locator items measuring any one specific skill. In vowel sounds, for example, six skills are being measured (short vowel sounds, long vowel sounds, vowel digraphs, vowel diphthongs, schwa, and vowel controllers), and therefore there are twelve Skill Locator items upon which to base judgments concerning performance on vowels.

The system was designed to be flexible. It can be made relevant to any reading program by adding new objectives or deleting undesired objectives. In the planning stages is a catalog of objectives from which a school system can choose so that Mini-Tests can be provided which exactly fit its program.

An experimental version of the Skills Monitoring System is currently being tried out in Grade 4 in three school systems in Michigan. Materials for Grades 3 and 5 are now being prepared. It was seen from the pilot testing of the system that the teacher needed something to help remedy problems indicated by the Skill Locator and Mini-Tests. Therefore a teacher support system is being built. This includes a sequence of workshops to be conducted at the local level, sample lesson plans, and a guide to instructional materials which will tell the teacher for each grade level for each objective where she can find appropriate teaching materials. This is the third and final major

component of the Skills Monitoring System.

The Skills Monitoring System with its Skill Locator, Skill Mini-Tests, and instructional support system seems to fulfill a need that has long existed in classroom situations insofar as both teacher and pupil are concerned. The materials provided will help to pinpoint individual and class needs, organize instruction, and follow each pupil's progress in attaining the skills of reading.

CLASSROOM MANAGEMENT CONSIDERATIONS
PROCESS AND PRODUCT PERSPECTIVES

G. Richard Zubulake, Ph.D.
Director of Instruction and Media
Clare County Intermediate School District
Clare, Michigan

The purpose of this paper is to present information on aspects of the local educational agencies' process management of the Harcourt Brace Jovanovich, Inc. Skills Monitoring System in Reading. Process evaluation is here defined as: (1) Examining the success with which the assessment materials are implemented. (2) Examining the instructional strategies employed following use of the reading assessment materials. Topics of the paper will range from reasons for entering into the program to a summary of specific process considerations.

Evaluation, for a number of years, has emphasized the selective function. Essentially, this task related evaluation to the identification of the few who were able to complete the academic program and be admitted to higher education. In contrast to using educational evaluation for selection purposes, current thought places its primary function as the development of the individual.

A number of current developments have made a systems approach to reading assessment timely. The most important of these has been the incorporation of the assessment on student learning as a basic part of the teaching-learning act. A second development has been the advocacy of criterion-based instruction. A third emphasis has been the advocacy of accountability, which relates to reporting student accomplishments in terms of valid measures on the objective units which each has worked.

Additionally, we at the local educational agency view innovation as essential if our schools are to be truly effective in achieving their desired ends. Continuing assessment of the process and the resultant product of our schools is also necessary. This, we believe, means the development of the principles and techniques for critically judging the worth of whatever the school teaches and the effectiveness and efficiency of our methods of instruction.

The approach incorporated in the HBJ Skills Monitoring System - Reading encompasses this philosophy and was our primary rationale for entering into a cooperative venture with Harcourt on field-testing the program.

Specifically our involvement is to field-test the HBJ Skills Monitoring System - Reading in three selected rural districts. Our further commitment is to aid in the evaluation of the reading assessment package as well as the instructional implications relating to it. In our portion of the project we have four fourth grade teachers from each of the three participating districts.

Preplanning was considered an integral component of our local educational agencies' involvement so that all personnel would clearly understand operational procedures, monitoring schedules, general goals, and time lines. This preplanning is evidenced by the comprehensive program evaluation report provided in Appendix A.

In the HBJ Skills Monitoring System, evaluation is viewed as a continuing procedure. We, as a participating agency, extend this assessment to

process considerations of the program with the hope that it will supply our local educational agencies with pertinent information which may be utilized in the continued improvement of their reading instructional programs.

First to be discussed are the aspects of process evaluation and the conclusions pertaining to it at this point in time. Process evaluation, in this project, is a primary concern. It deals particularly with the examination of the project methodology which is being implemented in each of our three participating districts and each cluster of four teachers. Included in this evaluation is the examination of the degree to which key personnel fulfill the prescribed assessment and instructional roles.

In our process evaluation, staff are periodically asked to respond to questionnaires and rating instruments which request information on, or about, the operational assessment and instructional elements of the project. Specific examples follow: (Appendix B)

I. District Profiles - Rating instruments which are used to assess, in detail, the degree to which the program is operational.

   A. Methodology

      1. Quality of instruction.
      2. Understanding of instruction.
      3. Time available for instruction.
      4. Learning materials.

B. Interaction patterns

C. Records

II. Observation - Systematic observation of the instructional setting and operations of the project occur at pre-set times by the intermediate coordinator. This information augments other data obtained from the aforementioned items.

III. Assessment commentary sheets. Sheets which may be utilized to summarize information pertaining to identified units of assessment.

The quality of assessment and follow-up instruction is viewed in relation to presentation, explanation, and ordering of elements in the assessment package which thereby determines the instructional focus. Our teachers generally rate the reading assessment as good since it approaches the optimum for each learner and at the same time allows a parallel approach to instruction. However, it is our feeling from follow-up instruction that we must find, or implement, more group instructional modes, since it is unlikely that schools will be able to provide instruction for each separate learner. Tutorial assistance or the one-to-one relationship represents the most costly type of help, and, we feel, should only be used when alternative procedures are not effective. Our experience to date, indicates that small groups (2-3 students) meeting regularly, were most effective. This was especially true when these students could help each other without any danger of

special advantages in a competitive situation. We, therefore, feel that post learning following the use of the HBJ Skills Monitoring System can be most effective when it is turned into a cooperative process involving the small group.

We have observed no problem in the understanding of assessment instructions. However, some difficulty has occurred in understanding learning instruction following assessment. This is defined as the ability on the part of the learner to understand the nature of the task which he is to learn and the procedure he is to use in accomplishing that task. Our attempt in the program has been to modify instruction to meet the needs of individual students. Another approach to differences in students ability to understand instruction is to vary the instructional material. I will not extensively comment here since my colleague's presentation centers around this and other relevant areas. Suffice it to say, that in the use of alternative methods and materials for instruction, our constant concern was that these be attempts to improve the quality of instruction as related to the ability of each student to understand the instruction.

The time allowed for evaluation in the reading assessment package poses no problems. In contrast, the time allowed for learning as a follow-up to assessment is a key. It has been stated that the time available for instruction should be made consistent with the needs of each student. This implies that the student must be allowed time for the learning to take place. Our view at this point in the program,

plus some other relevant reviewed research, indicates that it is just not the amount of time spent in learning that accounts for learning accomplishment. We feel that each student should be allowed the time he needs to learn a subject, but this also poses certain constraints. The task for follow-up instruction is, therefore, to find ways of altering the time individual students need for learning, as well as ways of providing whatever time may be needed. Thus, it not only relates to bettering the instructional program, but also includes the question of school organization, i.e., extended day, year around school, different grouping patterns, and materials considerations.

At this point in time our conclusions are tentative; but we feel that the HBJ Skills Monitoring System - Reading is one of the answers to better educational programming. However, the key still depends upon the teacher in the educational setting.

APPENDIX A

P.E.R.T. CHART

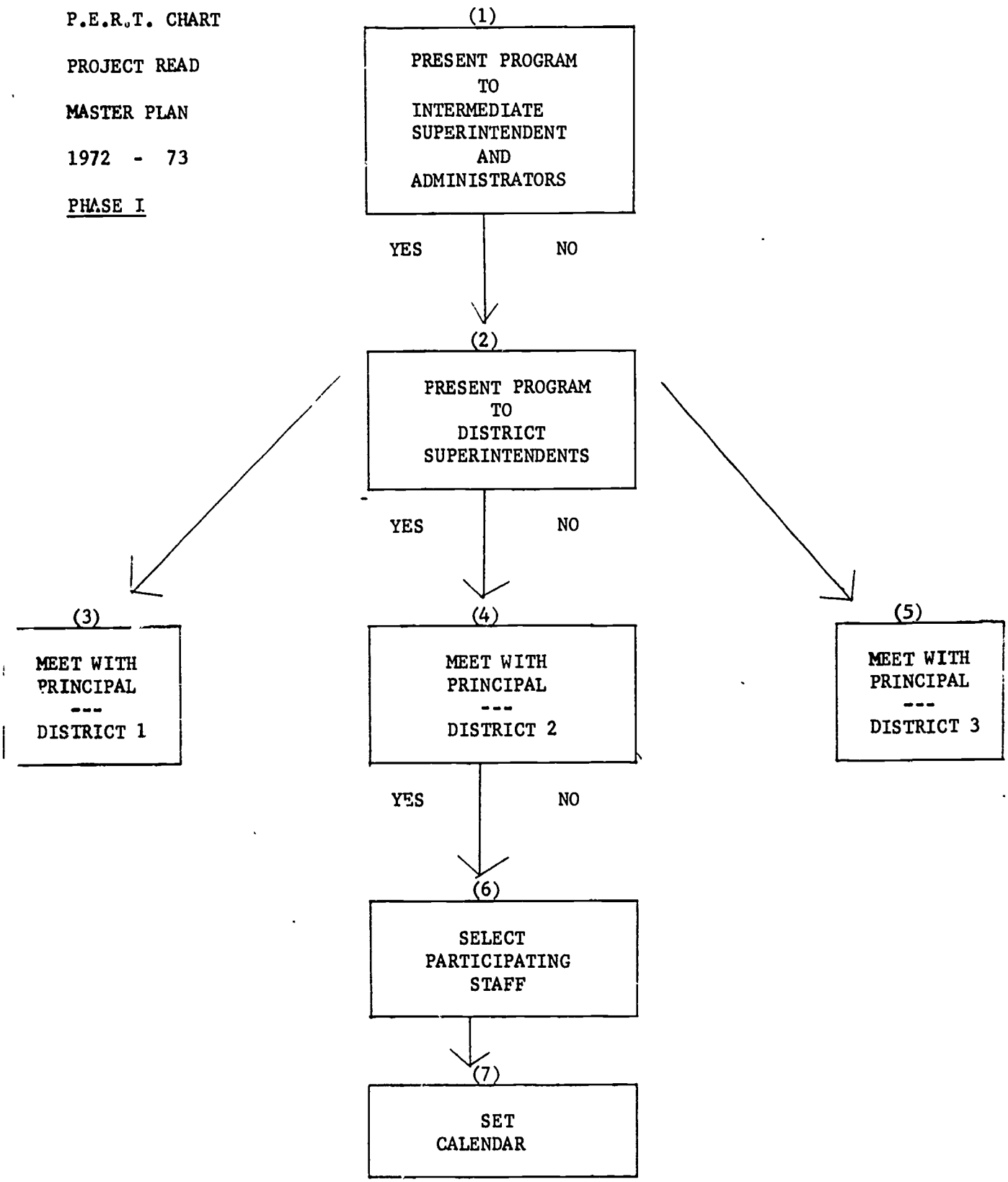
P.E.R.T. CHART

PROJECT READ

MASTER PLAN

1972 - 73

PHASE I



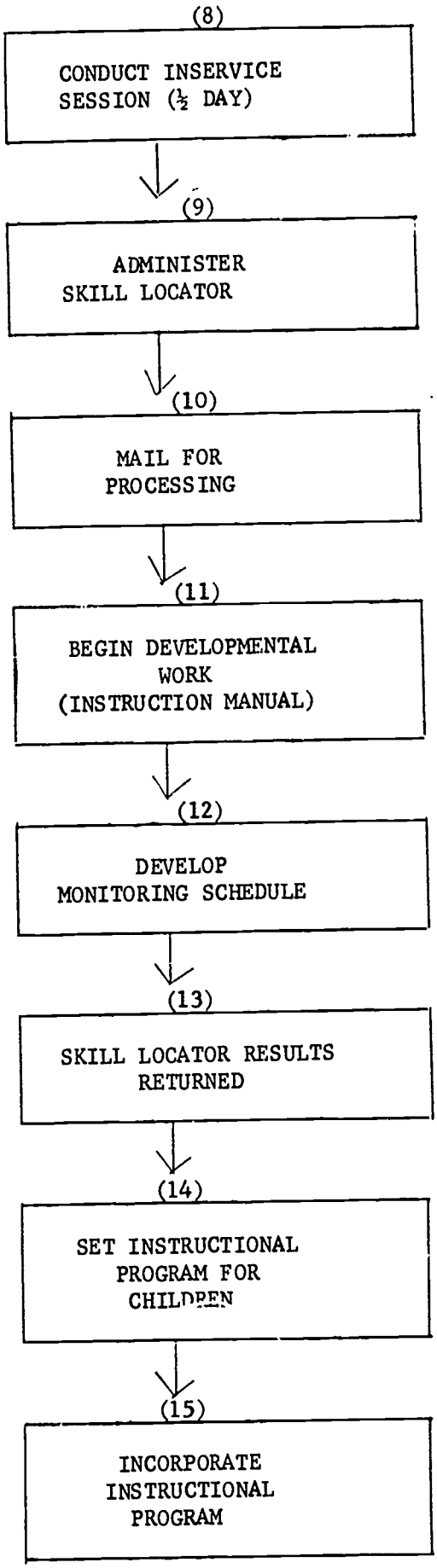
P.E.R.T. CHART

PROJECT READ

MASTER PLAN

1972 - 73

PHASE II



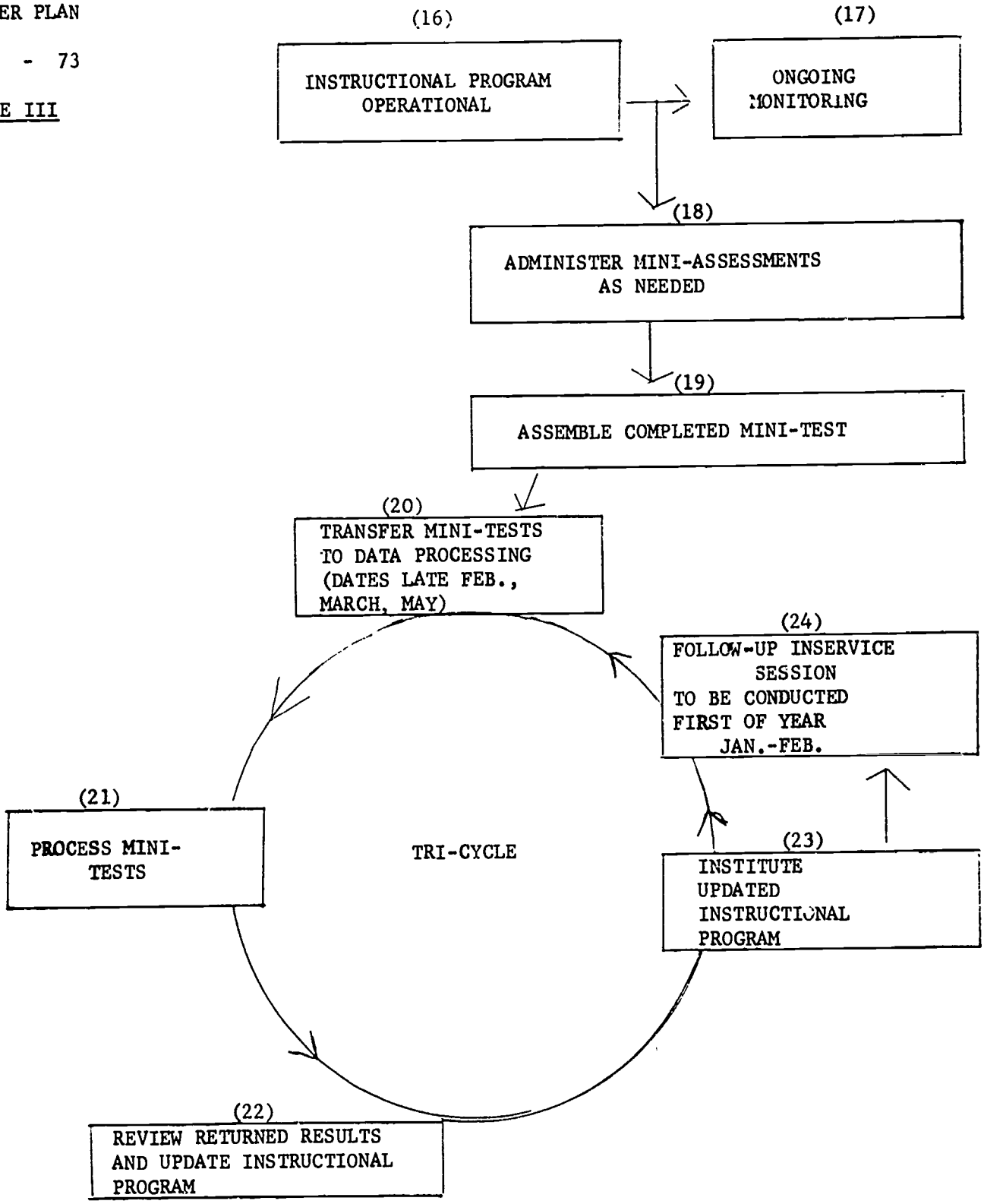
P.E.R.T. CHART

PROJECT READ

MASTER PLAN

1972 - 73

PHASE III



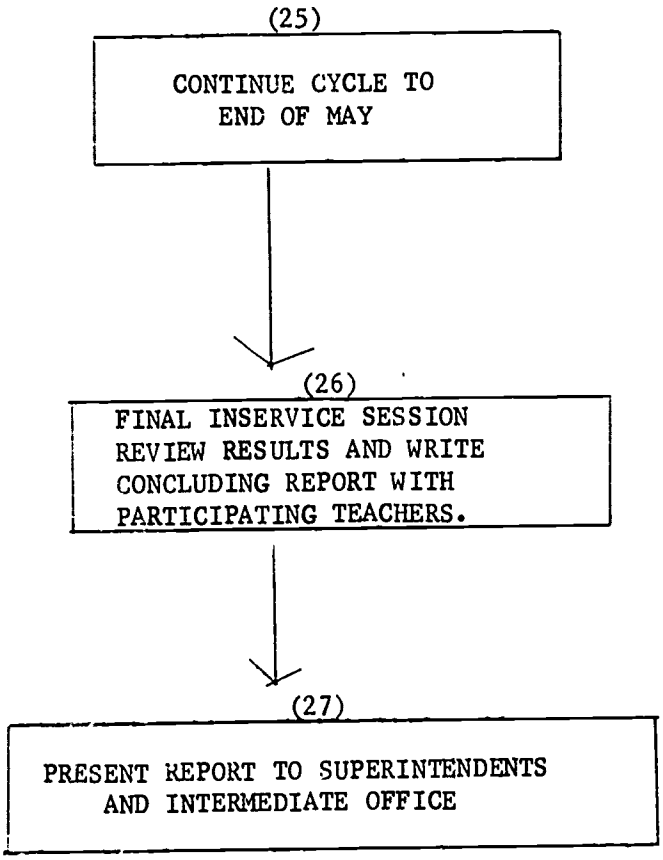
P.E.R.T. CHART

PROJECT READ

MASTER PLAN

1972 - 73

PHASE IV



APPENDIX B

PROJECT REACTION SHEET

PROJECT REACTION SHEET

Section I: Rank the degree of implementation of each element listed below. Circle the appropriate ranking that reflects your effort at this.

1 - No effort 2 - 25% effort 3 - 50% effort

4 - 75% effort 5 - 100% effort

	RANKINGS				
A. METHODOLOGY					
1. Lesson plans are stated in behavioral objective terms.	1	2	3	4	5
2. Locator tests are used for diagnosis.	1	2	3	4	5
3. Mini tests are used for diagnosis	1	2	3	4	5
4. Test results are analyzed	1	2	3	4	5
5. Appropriate prescriptions are used	1	2	3	4	5
B. APPLICATION					
1. Learning sequences are systematically outlined	1	2	3	4	5
2. Lesson plans are written	1	2	3	4	5
3. Individualized instruction occurs	1	2	3	4	5
4. Pupils are given opportunity to progress at their own rate	1	2	3	4	5

SECTION II. Rank the degree to which interaction patterns are operational.

1 - Never operational 2 - 25% operational 3 - 50% operational

4 - 75% operational 5 - 100% operational

		RANKINGS				
A.	INTERACTION PATTERNS (Teacher-Pupil Work)					
	1. Large groups exist	1	2	3	4	5
	2. Small groups exist	1	2	3	4	5
	3. Individual Pupil Activity exists	1	2	3	4	5
	4. Teacher assists individual pupils.	1	2	3	4	5
	5. Pupils request help.	1	2	3	4	5
	6. Discussion occurs between teachers and pupils	1	2	3	4	5
	7. Pupils contribute to group work	1	2	3	4	5
	8. Pupils are grouped and re-grouped as needed	1	2	3	4	5
	9. Evidences of teacher-pupil planning exist	1	2	3	4	5
	10. Pupils participate in self-evaluation, progress is shared with child.	1	2	3	4	5
	11. Opportunities for pupil-initiated work exists	1	2	3	4	5
B.	PUPIL ACTION-PEER INTERACTION					
	1. Pupils work with software	1	2	3	4	5
	2. Pupils use hardware	1	2	3	4	5
	3. Pupils check own work independently	1	2	3	4	5
	4. Pupils work independently with other pupils	1	2	3	4	5

SECTION III: Records: Rank each item according to the scale given. Circle the number.

1 - Infrequently done 2 - 25% current 3 - 50% current

4 - 75% current 5 - 100% current

A. SYSTEM-REQUIRED RECORDS:

1. Scores from regular testing	1	2	3	4	5
2. Health records	1	2	3	4	5
3. Anecdotal records	1	2	3	4	5
4. Other pertinent information.	1	2	3	4	5
5. Indication that outside referrals are made.	1	2	3	4	5

B. PROJECT RECORDS:

1. Scores from regular testing	1	2	3	4	5
2. Optional test records	1	2	3	4	5
3. Diagnostic information from Locator Tests and Mini Tests.	1	2	3	4	5
4. Pupil profiles and/or check-lists.	1	2	3	4	5
5. Records of mastery on Locator Tests and Mini Tests	1	2	3	4	5
6. Pupil-Family data records.	1	2	3	4	5

Section IV. Materials: Circle the number for each of the following. Consider only reading and language arts teaching.

0 - Not Available 1 - Available but Never Used 2 - Available but Seldom Used

3 - Available and Sometimes Used 4 - Frequently Used

5 - Very Frequently Used

A. TEXTBOOKS AND SOFTWARE

1.	Scott Foresman Basal Readers	0	1	2	3	4	5
2.	Scott Foresman Open Highway Readers.	0	1	2	3	4	5
3.	High interest, low difficulty books.	0	1	2	3	4	5
4.	Vocabulary flash cards.	0	1	2	3	4	5
5.	Word analysis activities.	0	1	2	3	4	5
6.	Comprehension workbooks or practice exercises.	0	1	2	3	4	5
7.	Barnell Loft workbooks.	0	1	2	3	4	5
8.	SRA Reading Laboratory.	0	1	2	3	4	5
9.	Other Basal Readers.	0	1	2	3	4	5
10.	Study skills practice exercises.	0	1	2	3	4	5
11.	Teacher produced or prepared materials.						
	a. Vocabulary Development	0	1	2	3	4	5
	b. Word Identification	0	1	2	3	4	5
	c. Comprehension	0	1	2	3	4	5
	d. Study skills	0	1	2	3	4	5

12. List other textbooks and software that you use frequently.

B. HARDWARE

1. Audio Flash Cards	0	1	2	3	4	5
2. Tape Recorders	0	1	2	3	4	5
3. Listening post	0	1	2	3	4	5
4. Movie projector	0	1	2	3	4	5
5. Filmstrip/slide projector	0	1	2	3	4	5
6. Overhead projector	0	1	2	3	4	5

7. List hardware programs (e.g. Systems 80) that you use frequently.

Section V: General comments (Optional):

APPENDIX C

READING ASSESSMENT

COMMENTARY SHEETS

SAMPLE

READING ASSESSMENT

COMMENTARY SHEETS

VOCABULARY DEVELOPMENT SERIES A	VOCABULARY DEVELOPMENT SERIES B
1.	1.
2.	2.
3.	3.
4.	4.
5.	5.

GENERAL COMMENTS

USE OF THE HBJ SKILLS MONITORING SYSTEM IN READING FOR INSTRUCTIONAL PURPOSES

Kenneth L. Carlson, Ph.D.
Coordinator of Language Arts and Reading
Wayne-Westland Community Schools
Wayne, Michigan 48184

Introduction

The age of accountability in education has arrived. Throughout the United States state departments of education, regional boards of education and local school districts have begun to develop accountability models for curriculum and instruction. The development of such models have, of necessity, been rooted in performance or behavioral objectives that can be used as a basis for evaluation of instructional competencies.

In the Wayne-Westland Community Schools (Wayne, Michigan) we are using the HBJ Skills Monitoring System in Reading to help our staff implement a plan of "instructional management by objectives." This management system is the basis for the establishment of a continuous progress program in reading and language arts that is targeted for completion by 1975.

Development and Use of the Assessment System

When the "instructional management by objectives" system was initiated last year, the prototype for diagnostic and evaluative teaching was based on the item analysis report of the Metropolitan Achievement Test. While this report was helpful in assisting classroom teachers in planning

instruction initially, it was woefully inadequate in assisting teachers with their planning for below average and above average readers. We knew that we needed a criterion referenced reading assessment system and began to plan the development of such a system. We began working with Dr. Muriel Abbott and Dr. Barrie Wellens on the prototype of the HBJ reading assessment system.

After the prototype for the HBJ system was developed, we field tested it in the Wayne-Westland Community Schools 1972 Summer School Program. Fourth and fifth grade students were given the HBJ Skill Locators. The data on the Skill Locators were compared with the individual diagnostic reading profile that had been given to each student earlier in the school year. The HBJ Skill Locator information and the individual reading profile data were correlated to form a prescriptive matrix for collating existing reading materials to the performance objectives. In this way, we were able to begin to develop a resource guide of instructional materials that could be used by classroom teachers in planning their individualized program.

In the fall of the year, twenty-two fourth grade teachers in the Wayne-Westland Community Schools volunteered to take part in the year long field test of the HBJ Skills Monitoring System in Reading. This group of teachers was made up of both self-contained classroom teachers and open space classroom teachers. All teachers in the project are attempting to establish a continuous progress program in reading instruction and were quite eager to have the HBJ system available to them.

The Skill Locators were administered in the latter part of October and beginning of November. Anecdotal records of reading behavior were logged while the Skill Locators were being scored and the computer printout was being prepared. When the teachers received the printout, they began to administer the Form A Skill Mini-Tests to refine the initial prescription. Adjustments in teaching patterns and use of instructional materials were also made. We are currently coding instructional materials used in the Wayne-Westland Community Schools to the performance objectives in the HBJ Skills Monitoring System in Reading. This process will continue for the remainder of the field testing.

Impact on Classroom Practice

In the Wayne-Westland Community Schools the adopted reading series is the 1965 edition of the New Basic Readers published by the Scott Foresman Company. The adoption includes the use of the Open Highways readers for below grade level students. The testing program which accompanies the Scott Foresman reading program is also used to evaluate the progress of developmental reading instruction. Up-to-date reading records are maintained so it is possible to review the reading progress of students at fairly regular intervals.

When the "instructional management by objectives" system was used during the 1971-72 school year, elementary teachers continued to rely heavily on the Scott Foresman reading materials for the majority of their teaching activities. However, as the teachers began to grow in their understanding

of individual instructional needs, they began to search for supplementary materials that could be used for meeting these needs. During the latter part of the 1971-72 school year and the beginning of this school year, a wide variety of teaching materials were purchased. In this way the resources available to classroom teachers increased immensely.

The teachers using the HBJ Skills Monitoring System are beginning to develop proficiency in two broad areas: (1) the use of flexible grouping for instruction and (2) the use of multi-level skill development materials. We anticipate that the teachers will continue to exhibit growth in these areas as they continue to use the HBJ System. Even in the short time that the teachers have been using the HBJ Mini-Tests, they report that changes in instructional programming are made more frequently.

In a recent survey of the twenty-two Wayne-Westland teachers using the HBJ System, the teachers stated that they were able to plan more efficiently for reading skill instruction. They claim to be able to use commercial materials in a more prescriptive fashion. The teachers were asked to rank their use of various instructional materials. They rated their use of reading materials according to the following criteria:

0 - Not available

1 - Available but never used

2 - Available but seldom used

3 - Available and sometimes used

4 - Frequently used

5 - Very frequently used

This survey did indicate that two-thirds of the pilot project teachers still use the Scott Foresman basal readers very frequently for reading instruction. Two-thirds of the teachers also reported a moderate to very frequent use of high interest, low difficulty books for the directed teaching of reading. While other basal readers are readily available, the project teachers stated that such materials are moderately or seldom used for reading instruction.

As the teachers develop increased proficiency in flexible skill grouping for reading instruction, they also tend to select other materials for specific skill instruction. All teachers report a moderate to very frequest use of vocabulary flash cards, word analysis activities, comprehension workbooks and comprehension and study skill practice exercises. Materials that require the use of hardware such as tape recorders, listening posts or filmstrips are being used with increasing frequency for specific skill instruction.

One of the most notable aspects of the survey revealed that all teachers reported a moderate to very frequent use of teacher prepared or produced materials in the four skill areas in the HBJ Skills Monitoring System: vocabulary development, word identification, comprehension and study skills.

Conclusion

While the HBJ Skills Monitoring System in Reading has been in use in Wayne-Westland for a few short months, its impact has been a substantial one. The teachers working in the pilot project have improved their ability in the diagnostic-prescriptive teaching of reading. The HBJ

Skills Monitoring System in Reading is helping our school district reach its goal of continuous progress programming in reading instruction.

THE IMPACT OF CRITERION-REFERENCED READING ASSESSMENT ON PLANNING, TEACHING AND EVALUATING IN A RECENTLY DESEGREGATED URBAN SETTING.

Richard C. Benjamin, Ph.D.
Director, Office of Evaluation Services
Lansing School District

The focus of this paper is the impact of a pilot project in objective-referenced reading assessment on two organizational goals of the Lansing School District. One goal is the classroom utilization of a comprehensive evaluation model. The other goal is the implementation of a desegregation plan. The intent of this preliminary investigation was to determine how reasonable it is to depend to a greater degree on objective-referenced measurement to further the above organizational goals.

An experimental design was not employed for two reasons. First, the size and timing of the current project did not warrant the sophisticated instrumentation that would have been necessary. Second, many teachers are already utilizing objective-referenced assessment systems as fast as they are available and will continue to do so regardless of documented relationship to a system evaluation model or desegregation plan.

The Evaluation Model

A comprehensive evaluation model is being implemented in the Lansing School District. The Context-Input-Process-Product (CIPP) Model advanced by Stufflebeam, was selected because the primary function of the model is to assist decision-making in a decentralized setting. Our particular decentralization plan, called responsible-autonomy, does not involve

mini or regional boards of education. Rather, it focuses on school level interaction of parents, students, faculty, and administration. Their joint task will eventually be to address the major functions of planning, hiring, implementing, and evaluating their school program. The constraints would be the same as those operating centrally at the present time; legal, financial, and those educational goals appropriate for the district as a whole.

It was evident, that as we moved from a relatively centralized to a relatively decentralized operation such as described above, local school building personnel must acquire specific skills in the gathering and analyzing of data appropriate to the decisions they will be making. Of the existing evaluation strategies, the CIPP model seemed the most appropriate.

This model essentially involves four types of activities. Context evaluation is determining what needs are most important and the constraints operating in meeting the needs. Input evaluation consists of a systematic consideration of various strategies to meet the priority needs and the selection of the most appropriate strategy. Process evaluation entails monitoring of the implementation to insure that it conforms with the planned strategy. Product evaluation is finding out if the implemented strategy is producing the desired results in the short and long term. One focus of this investigation is the impact of the pilot reading assessment program on local building personnel with regard to their willingness and ability to engage in the major activities of the evaluation model.

The Desegregation Plan

In Lansing, a Board of Education policy on equal educational opportunity has existed since 1964. That policy asserted that ethnically integrated education was a higher priority than neighborhood schools. Until 1972 this affected only inner-city black students who were bused to predominantly white schools. In September 1972, the first phase of the comprehensive desegregation plan for elementary schools was implemented. The total plan was to extend over a four year period. In phase one, eight schools were divided into two clusters of four buildings each. Within each cluster, two-way busing was initiated for middle and upper elementary students from predominantly white neighborhoods. These children were bused to schools previously serving mostly minority students as were minority students sent to schools previously serving mostly white students. The plan provided for the addition of a third cluster of five buildings during year two, evaluation of results and population shifts during year three and total district desegregation during year four.

Previous to the above plan, approximately 500 minority youngsters were transported to serve the policy on equal educational opportunity. Around 15 majority youngsters from the same neighborhood were included in that one-way effort. Phase one of the above plan, involving two-way busing, provided for busing 290 minority and 940 majority youngsters for an average ride of 13 minutes each way. The longest run takes about 20 minutes each way.

More important to our purposes, the plan dealt not only with the

transportation of students to reduce minority isolation but also had a strong instructional component. Two goals were identified for the students involved; improved academic achievement over previous years; and improved attitudes towards self, others, and school. Three processes were identified to assist in accomplishing the above objectives: individualized instruction, increased staff training opportunities, and community involvement. It must be remembered that these objectives and processes were part of a four-year plan with the third year intentionally set aside with no new busing activities specifically to allow time for proper measures to be developed, pretesting to be analyzed, and progress to be measured over a longer period than is possible in most projects.

The board members voting for the above plan (5 out of 9) have since been recalled and replaced. The new majority of the board rescinded the entire plan, with the first phase to be terminated at the end of this school year. Their action is, of course, being challenged in the courts, and should the outcome be that the original plan will be implemented, it will still be important to determine the extent to which the pilot assessment program reinforces the objectives and processes of the desegregation program.

The Methods

The pilot reading assessment plan was made available to teachers on a voluntary basis. Extra attempts were made to assure that teachers in the cluster schools were aware of it. It was not considered critical to place the program primarily in cluster schools since there were plenty of non-cluster classrooms that were desegregated by virtue of their

neighborhood composition or because of the one-way busing which would continue until the fourth year of the plan. Since, at the time, the entire system was on a course leading to implementing both the program of responsible-autonomy and the desegregation program (including the educational component) the impact of the pilot program on the two could be assessed, at least informally, inside and outside the cluster. The successful recall weakened this assumption because many principals and teachers could, with some justification, question the sense of exerting the necessary energy to make a doomed plan work. In spite of this, we decided to go ahead with a severely curtailed design that would at least give us direction for future activities, should they be possible.

The most serious consequence for the staff involved may have actually helped to remove an alternative hypothesis. If individualized instruction and parent participation were no longer to be part of a Board sanctioned plan, increases in these areas might be more likely the result of the introduction of the pilot program.

We ended up with 35 teachers participating in the project by using one of four commercially available objective-referenced assessment systems. Harcourt Brace Jovanovich, Inc. Skills Monitoring System in Reading was used by 23 teachers, CTB-McGraw Hill Prescriptive Reading Inventory was used by four teachers, Educational and Industrial Testing Service materials were used by five teachers, and Random House Criterion Reading materials were used by three teachers. Obviously, the results reported will heavily reflect the impact of the HBJ pilot assessment program because

of the disproportionate number of teachers to whom they made materials available.

Given the above problems, highly structured measures of the dependent variables were not possible. Therefore, tight investigation of the extent to which school personnel were implementing the desegregation plan and planning ways to improve it the following year, was not appropriate. Instead, we went ahead with structured interviews, questionnaires, classroom observations, observation of parent meetings, and inspection of pupil records. We were forced to depend upon participating teachers and principals to make the connection between the pilot assessment program and the key dimensions of the evaluation plan and the desegregation plan.

The Results

The impact of the pilot assessment project on the four activities of the evaluation model was probably the most obvious. Many of the teachers made remarks and evidenced behaviors that clearly attributed increased activity in three of the four areas to their involvement in the pilot program. Examples in each of the areas follow.

Context Evaluation

Almost without exception, teachers indicated that they knew more about the reading skill needs of their class as a whole and of each individual student than previously. This is perhaps a self-evident finding but in many cases we are speaking of teachers with considerable experience. Several comments were made concerning the ease with which teaching

priorities could now be established and how also, the constraints of the closed classroom were more evident.

Input Evaluation

Implementation of the pilot program led to sophisticated consideration of various teaching plans. The comment that the objectives and assessment materials should be tied to specific instructional treatments was second only to comments about how late the results from the publisher were. Principals were aware that teachers were increasingly focused on selecting instructional strategies appropriate to identified needs rather than focusing on the problems of implementing a given strategy uniformly with most pupils.

Process Evaluation

Due to the anticipated problems with the production of testing materials and data processing, results from the publisher were held up much longer than will be the case when the program is operational. This probably explains the absence of any remarks or observations in this area.

Product Evaluation

As expected, the impact in this area was noticeable in several ways. Pupil records were often proudly displayed which showed careful monitoring of pupil progress on specific skills during very short time periods. Most participants indicated in one way or another a much firmer idea of not only where students were, but their rate of progress.

The Desegregation Program

The impact of the pilot assessment project on the major activities within the desegregation program was generally more difficult to determine. This was probably due to the on-again off-again status of the desegregation effort in the past few months. In spite of this situation, some clear findings were obtained.

Pupil Objectives

The most frequent comments in these areas had to do with the fact that short term cognitive growth of each pupil was documented. The impact of this on the amount of achievement was noted in several cases. Students were doing more and better in a situation where success was possible and likely in the short run. One teacher admitted that he honestly saw all students as learners for the first time because he had evidence that they were indeed learners. The reinforcing effect of his new attitude could well be one of the most significant outcomes of the project.

With regard to attitudinal growth of the students, indications were that the pilot assessment program had an impact on attitudes toward self and attitudes toward school if it had any effect at all in this area. However, no teacher reported any connection between the pilot program and growth in student attitudes towards others. This was to be expected in the short run. Hopefully however, as better attitudinal measures are employed and as the program is operational over a longer time, results in this area should be obtained.

Process Objectives

The introduction of the objective-referenced reading assessment system was hypothesized to have the most direct bearing on the establishment of a more personalized learning setting. Indications are that this was indeed its major area of impact.

The specific measures of this objective center around a few core dimensions of individualized instruction. What we were looking for was evidence of the following:

- the teacher is employing some sequence of performance objectives, at least to the extent that the instructional treatment for a given student at a point in time is partially explained in terms of what he has recently mastered, what he is currently addressing, and the next objective in the sequence.

- the teacher has provided a setting where the student has a more than even chance of experiencing success at the task at hand, whether it is primarily a teaching or a testing activity.

- where appropriate, alternative methods of instruction or learning are available to reach the same objective, and that the methods are determined to some extent for each child taking into account his interests and/or learning style.

We have no data relating to the important questions concerning interaction of various teacher variables with the effects resulting, but we have a clear reading that participating teachers previously not providing settings like the one described above now do, and personally attribute the change to the availability of sensible objectives and related test items.

The largest impact was on the assignment of pupils to instructional treatments based on need. Some change was noted with regard to the measures of increased pupil success and responsiveness to learning style. Further progress here seems to await greater teacher awareness of a wide range of instructional treatments to accommodate more levels of success and diverse pupil interests and learning styles.

In addition to the expected impact on the instructional setting, other very clear effects resulted with regard to the other two process objectives of the desegregation plan, parent involvement and staff training. The most interesting to program planners was the tightening of communication concerning pupil progress between teacher and parent. It was frequently mentioned by teachers that conferences and conversations became less vague and more emphasis was placed on pupil success. Areas of potential parent participation were more easily identified. Teachers previously had attempted to steer parents away from focusing on the achievement of their children relative to other children and instead recognize the specific accomplishments that were demonstrably related to a sequence of objectives. Normative data was finally interpreted in terms of some recognizable tasks.

The pilot program also impacted the staff training effort. Teachers became aware of gaps in their own expertise that they now saw as quite important. Needs have been identified in the areas of knowledge of new instructional materials, diagnosis of interests and learning styles, and knowledge of methods [illegible]aching specific reading skills.

Conclusions

Given the information collected it could be decided that the pilot assessment program did reinforce the main objectives of the evaluation model and the desegregation plan. No comments were made by teachers and no observations were made that indicated the pilot assessment program ever detracted from, either the activities of the evaluation model or the objectives of the desegregation plan.

There was a tangentially related concern that the individualization of instruction was antithetical to the objectives of improved self-concept and improved academic achievement within the desegregation program. The fear was that minority student self-perceptions of academic inadequacy would be reinforced and that the social setting of the classroom would not be fully utilized to communicate high teacher expectations for each pupil. Preliminary findings contradict this concern.

Other than the above, the information was uniformly supportive of objective-referenced measurement, especially with regard to Context, Input, and Product evaluation activities; and with regard to individualization of instruction and parent participation within the desegregation plan.

The responsiveness of the pilot assessment system to the unique demands of the desegregation plan is most encouraging and most frustrating, since community evaluation and decision-making in this area will clearly progress without regard for data of any sort.