THE DESIGN AND IMPLEMENTATION OF A CRITERION REFERENCED INSTRUCTIONAL SYSTEM: AN ALTERNATIVE TO INSTRUCTIONAL SERENDIPITY

by

Huntley Vaughan Hoffman

A Thesis Submitted to the Faculty of the

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF ARTS

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 7 4

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Huntley V. Hoffman

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

John R Bergan — 5/20/74

JOHN BERGAN — Date

Professor of Educational Psychology

This thesis is dedicated to

Sally Vaughan Hoffman

whose gentle and patient prodding is appreciated

ACKNOWLEDGMENTS

The author would like to thank both Barbara Tallis and Patti Specio for their enthusiastic and meaningful contributions to this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

ABSTRACT

There is a great deal of dissatisfaction with education today due to a failure to provide many students with the skills necessary to compete in a modern society. A significant contributor to this problem is instructional deficiency in the classroom. Too many teachers are unaware of the precise needs of their students and consequently cannot provide efficient instruction.

To alleviate this instructional deficit, teachers might incorporate a Criterion Referenced Instructional System (CRIS) into their teaching program. A CRIS is comprised of instructional objectives, evaluation of each student's needs in relationship to the objectives, and an on going student progress assessment.

A research study was designed to both explore the feasibility of teachers designing an "in house" CRIS and, if successfully completed, to evaluate how meritorious are the effects of using it in the classroom. Results of the study show that the classroom utilizing the CRIS had significantly higher reading scores than the control classroom, but no significant differences were demonstrated in other areas. The lack of more widespread success is attributed to the limited number of instructional hours included in the experimental period.

CHAPTER 1

THE PROBLEM

This paper was written to discuss the existence of an instructional deficiency in education, the etiology of the problem and a means to eradicate it. The paper is divided into four chapters. This chapter is concerned with the substantiation of the problem, the problem's origin, and the rationale for the research study presented in Chapter 3. Chapter 2 reviews the educational literature on pertinent related areas: diagnostic teaching, criterion-referenced tests, norm-referenced tests, and behavioral objective-based teaching systems. Chapter 3 presents a research study on the results of the implementation of a criterion referenced instructional system (CRIS). Chapter 4 is a discussion of the results of the study.

A Problem Exists

Signs of the existence of instructional deficiency are found in state legislatures, courts of law, and, of course, classrooms. The California State Legislature has enacted an educational law, the Stull Bill, which prescribes a uniform accountability system of evaluating teacher's performance according to student achievement (Neel 1972). School superintendents in Arizona are complying with a statewide directive for each district to generate behavioral objectives to be used to evaluate student progress. Another sign of educational deficiency

occurred in Oakland, California, where a high school graduate sued the school district in which he had been enrolled because it had failed to live up to its "teaching obligation." Classrooms present the clearest evidence of a learning problem. Throughout the country, schools are filled with students who are operating below grade level and accelerating at an intolerably depressed rate. If this instructional deficiency in our educational system is going to be alleviated, the contributing factors will have to be identified.

The Etiology of the Problem

There is no doubt that a student's academic progress is a function of his total environment and to place the entire blame for unactualized educational potential on instruction would be extremely narrow minded. However, education must "shoulder" the blame that can be attributed to the provision of less than optimal instruction. It should be noted that the word "instruction" not "educational environment" was used for the criterion. The rationale is that an optimal educational environment is fiscally untenable, but improved instructional procedure is relatively easily attained.

The most significant encumbrance to effective instruction is a teacher who is not aware of what a student needs to be taught in order to achieve a goal. Such information is vital to prevent a teacher from reteaching what a student already knows or trying to teach what a student is not ready for. Lack of effective techniques for both diagnosing a student's abilities and evaluation of his progress is a common weakness in the contemporary instructional process (Cheikin 1971). If a

teacher cannot evaluate a student's skill level and progress, then effective teaching is a matter of serendipity. There are numerous factors that perpetuate the unilateral concept of education (i.e., I know what I want you to do, but I don't know what you can do) that stifles educational productivity: norm-referenced test (NRT), educational specialists, means-orientated instruction, and lack of adequate teacher skills.

Norm-referenced tests are the main diagnostic tool used in contemporary education and they are inadequate for diagnosing. A diagnostic instrument should accurately measure student progress. These NRTs are capable of measuring only gross changes of learning and hence lack the sensitivity to evaluate the progress of many students (Block 1971; Brazziel 1971, Neel 1972, Tyler 1971). Also, a diagnostic instrument should have remediational value. Norm-referenced tests do not offer adequate information upon which to effectively formulate an educational program (Cummings 1972, Falcone 1969, Hawes 1973, Larsen 1972, Mann and Proger 1973, Prillaman and Prouty 1970, Ward 1970). The NRT scores are usually limited to translation into categories and stanines, rather than the specific strengths and deficits necessary to devise remediational plans.

Student achievement is also inhibited because many teachers feel that diagnosis is the domain of specialists only. This philosophy prevents teachers from actualizing their own diagnostic potential in the classroom (McCarthy 1971). The purpose here is not to de-emphasize the value of specialists for "special problems," but rather to

criticize school personnel who feel that diagnosis by both specialists and teachers is a mutually exclusive phenomenon.

The third detriment to effective instruction is an over emphasis on "means-referenced" instruction. The means-referenced concept of instruction is inadequate for most instructional decision making (Bloom 1968, Glaser 1968, Trueblood 1971). Emphasis on the "how" of instruction and neglecting the "what" does not afford pupils and teachers ample guidance and reinforcement to maximize learning. However, it is important to note that, as long as provisions are included to chart direction and monitor progress, instructional flexibility and experimentation is certainly an asset. Popham (1971) writes that numerous teaching techniques can be used to obtain the same educational goal.

The last cause of instructional deficiency to be discussed is that many teachers simply do not possess the requisite skills for more effective teaching. Many student difficulties are a function of teachers' inability to formulate relevant instruction based on valid, definable goals (Cheikin 1971). Popham's (1971) study confirms the existence of this condition. He found that experienced teachers were unable to promote attainment of prespecified instructional objectives any better than nonteachers.

Briefly, to be optimally effective a teacher must know what skills a student possesses, what goal the student is attempting to achieve, and if progress is being made. These evaluations necessitate that the teacher have certain skills: task analysis, ability to formulate behavioral objectives, design diagnostic instruments and

probes, and record keeping. Most teachers are not taught these skills in college, in-service, or workshops. If the causes of instructional deficiencies are to be eradicated, a pragmatic solution must be offered.

Purpose of Research Study

This study was designed to demonstrate two hypotheses. First, it is feasible for a teacher or group of teachers to design a CRIS which is keyed to their own educational objectives and teaching style. A CRIS is composed of behavioral objectives, a diagnostic instrument, progress probes, and a record keeping system. Second, if such a CRIS is incorporated into a teacher's instructional program, academic progress for that class will be significantly greater than for a classroom that doesn't utilize a CRIS.

The rationale for this study is to contribute to a body of knowledge that will facilitate actualization of teachers' instructional potential. Four occurrences will facilitate this result. First, teachers must be given a diagnostic and monitoring alternative to NRTs. Second, teachers must feel diagnosis is within their capabilities, not just for specialists. Third, teachers must have confidence in a CRIS approach to instruction. Fourth, teachers must have the skills to design and implement a CRIS. The author hopes that this paper contributes to these goals.

The next chapter discusses various author's opinions on subjects that are pertinent to the study. It should be noted that CRIS is

a term coined by the author to label his specific representation of diagnostic teaching. Since the term CRIS does not appear in the literature, it is not used in Chapter 2.

CHAPTER 2

REVIEW OF THE LITERATURE

This review of the literature is included to provide the reader a more meaningful context within which to evaluate the purpose and implications of this study. There are five topics discussed in this section. The first topic is a discussion of the general concept of diagnostic teaching. The subsequent topics--diagnostic instruments, objective-based instruction, task analysis, and assessment--are all component parts of diagnostic teaching.

Diagnostic Teaching

Diagnostic teaching is a term for an instructional philosophy, rather than a specific teaching method. The format of diagnostic teaching is built around an input schedule of information about the student. The following is an example of a diagnostic teaching format (Falcone 1969, Smith 1972): (1) assessment of needs of students, (2) instructional objectives based on those needs, (3) techniques and materials to achieve the objectives, (4) ways of observing effectiveness of teaching and learning, (5) adjustment according to the information that is collected.

Diagnostic teaching is the logical outgrowth of the recognition of individual differences. All children are unique at conception and become increasingly divergent with age and experience (Miller 1970).

This individual diversity makes the effectiveness of diagnostic teaching a function of the teacher's insight into each student (Bond 1972, Neel 1972). The remainder of this discussion of diagnostic teaching will consider the advantages and disadvantages of diagnostic teaching, as well as some current implementation of diagnostic teaching programs.

Advantages

Many authors feel that diagnostic teaching is extremely beneficial to students. Miller (1970, p. 420) writes, "Through diagnostic, individualized teaching, the gap between a child's innate creative talent and his actual output can be narrowed." Concurring, Falcone (1969, p. 44) states, "The model of the diagnostic classroom--or better the diagnostic teacher--breaks through the insidious psychomedical and educational stereotyping and strikes a heavy blow at present standardized psychological and educational tests as sources of relevant educational prescriptions for John and Mary." Other advantages of diagnostic teaching mentioned in the literature are: a reduced need for special education classes, decreased absenteeism, less vandalism, and increased student-parent enthusiasm about school.

Disadvantages

Some authors qualify their general enthusiasm about diagnostic teaching. Cronbach (1967) writes that teachers probably have a tendency to over differentiate between students when they diagnose. He cites his research on the inability of school counselors to predict student grade averages from standardized tests as part of his rationale. Cronbach

(1967, p. 30) summarizes his view on utilizing diagnosis in the classroom in these words: "Certain reasonable assumptions, entered into a decision-theoretic model, lead us to the conclusion that the poorer the differential information, the less the teacher should depart from the treatment that works best on the average. Modifying treatments too much produces a worse result than treating everyone alike." Another qualification to diagnostic teaching is that social skills are at least as important as academic skills; therefore, the virtue of using diagnostic teaching with children, especially socially deviant students, would depend on whether or not its use can improve social as well as academic functioning (Christoplos 1970).

Current Implementation

A widespread trend to diagnostic teaching is perceived by some contributors to the literature. Trueblood (1971) feels elementary mathematics teachers are discarding the traditional roles of "directors" and "lesson planners" and becoming "programmers" for individual learners. Massad (1972) states that she has seen diagnostic teaching used effectively in a multitude of reading programs with a diversity of instructional approaches. The starting place for any diagnostic teaching program is the assessment of student needs.

Diagnostic Instruments

Diagnostic teaching requires an objective method of evaluation to serve as a base. This is not intended to minimize the virtue of ongoing subjective monitoring as a means of evaluating social and

academic skills; however, subjective evaluation is simply too imprecise to allow for optimal academic progress.

An instrument that requires responding to questions with "right" and "wrong" answers is quite appropriate to meet this objectivity requirement. The two types of instruments that could be used for objective assessment are NRTs and CRTs. Both of these instruments will subsequently be discussed in regard to: definitions, differences, advantages, disadvantages, current status, and commercial availability.

Definition - NRT

The NRTs are tests which are used to evaluate a student's performance in relationship to the performance of other students on the same measuring device. The label norm-referenced originates from the fact that the individual is compared with some norm-referenced group (Popham and Husek 1969). Typical examples of norm-referenced measures would be standardized tests of achievement or intellectual ability.

Definition - CRT

A CRT is an instrument that is used to ascertain a student's ability relative to a criterion. Glaser (1971, p. 6) defines a CRT as a "test that is deliberately constructed to yield measurements that are directly interpretable in terms of specific performance standards." The term criterion-referenced is derived from the fact that the individual is compared to some criterion, rather than other individuals. It should be noted that a CRT is also known in the literature as "mastery testing" and "domain-referenced testing" (Mann and Proger 1973).

Differences between NRTs and CRTs

Basically the difference between NRTs and CRTs is that the latter are constructed to maximize the discriminations made between groups treated differently and to minimize the differences between the individuals in any one group, while NRTs are constructed to maximize the discriminations made among people having specified background and experience (Glaser 1963).

Simon (1969) writes that differentiating between CRTs and NRTs is illogical. His rationale is that since either instrument can generate absolute (criterion-referenced) or relative (norm-referenced) scores, the labels of criterion-referenced and norm-referenced tests should be replaced by norm-referenced and criterion-referenced scores.

Some examples may help illustrate the differences between CRTs and NRTs. If a degree of selectivity is required, then a NRT would probably be appropriate (Popham and Husek 1969). In other words, NRTs are optimal if a selection has to be limited to a prescribed number of individuals, i.e., promotions on an assembly line. In contrast, a CRT is appropriate to evaluate skill levels if there are no numerical constraint, i.e., students in a life-saving class trying to pass the final test. The divergent purposes of CRTs and NRTs are also manifested in differences of item construction, reliability, validity, item analysis, and score interpretation.

Item Construction. Item construction for CRTs and NRTs is differentiated by the purpose of the item writer (Glaser 1963, Popham and Husek 1969). The author of a NRT wants variability. The author of

CRT test is concerned primarily with accuracy of criterion measurement.

With the quest for variability paramount, NRT authors generate questions that will create a range. This necessitates the limitation of test items that either all examinees will correctly answer or virtually no examinees will answer. In addition, an author of a NRT may try to increase the allure of wrong answer options. The item difficulty manipulation utilized to generate variability usually results in the majority of questions on the final form of a standard test being answerable for 40 to 60 percent of the examinees (Tyler 1971). In contrast, the CRT item writer is not concerned with variability.

The emphasis for CRTs is that the item be an accurate reflection of the criterion behavior. It is of no moment whether the test items are difficult or easy, discriminating or indiscriminate. Those who write criterion-referenced items are primarily concerned with "defining the domain of relevant test responses" and the appropriate circumstances for demonstration (Popham and Husek 1969, p. 4).

Reliability. Once test items are constructed, reliability estimates can be used to refine the selection. Indices of reliability measure either internal consistency or temporal stability.

The reliability measurement for CRTs and NRTs is different. The inter-item reliability measure of a NRT is dependent on score variability (Harris 1972). In contrast, Popham and Husek (1969) have written that such typical indices of internal consistency and temporal stability are not appropriate for CRTs. Popham and Husek (1969, p. 3)

state that "a criterion-referenced test should not be faulted if, when administered after instruction, everyone obtained a perfect score. Yet, that would lead to a zero internal consistency, something measurement books don't recommend." Popham and Husek (1969) further state that criterion-referenced tests may well indeed have high inter-item and/or highest-retest correlation which is fine, but lack of such high correlations is not necessarily indicative of lack of consistency.

There are some suggestions in the literature to resolve reliability measurement for CRTs. Livingston (1972) suggests that concepts based on deviation from the mean will be replaced with a corresponding concept based on deviation from the criterion score. However, Harris (1972) concludes this theory is untenable for reasons beyond the statistical scope of this paper. Another resolution offered by Popham and Husek (1969) is that estimates similar to the standard internal consistency formulas be constructed to take larger temporal units into consideration. For example, pre- and postinstructional examinations could be developed that measure the ability of a test to produce variation from preinstruction to postinstruction testing. He adds that this could be a criterion for internal consistency, regardless of variability restrictions. To measure temporal stability of a CRT, Popham and Husek (1969) feel that it may be plausible to develop a confidence internal around the individual score.

Validity. Procedures for determining NRT validity, as with reliability are normally based on score variability. However, Popham and Husek (1969, p. 5) state:

> Criterion-referenced measures are validated primarily in terms of the adequacy with which they represent the criterion. Therefore, content validity approaches are more suited to such tests. A carefully made judgment based on the test's apparent relevance to the behaviors legitimately inferable from those delimited by the criterion is a general procedure for validating criterion-referenced measures. Construct validity may be utilized for either CRTs or NRTs to support confidence in the instrument.

Item analysis. Another difference between NRTs and CRTs exists in the means of item analysis. For NRTs item analysis procedures are utilized to identify and delete indiscriminate test items. Typical examples would be items that are too easy, too hard, and/or ambiguous. With CRTs nondiscriminative items may be retained on the test. An indiscriminate item on a CRT should remain if it reflects an important attribute of the criterion (Popham and Husek 1969).

Positively and negatively discriminating items are treated in the same way for both NRTs and CRTs. Positively discriminating items are retained on both types of tests. For CRTs, positively discriminating items may even be an aid to analysis of where the training program is ineffective (Popham and Husek 1969). Test items which are more apt to be answered by low scorers than high scorers (negative discriminators) are deleted on both CRTs and NRTs. However, Popham and Husek (1969) advise that the flaw of negative discriminators in CRTs can usually be discovered.

Score Interpretation. Another index of distinction between CRTs and NRTs is the interpretation of test scores (Simon 1969, Livingston 1972). The NRT scores are usually translated into percentile rank or group scores in order to analyze how an individual performed in relation to the group. In contrast, CRT scores are usually

interpreted by their position above or below a fixed standard (Ebel 1971). In other words, criterion was met or it wasn't. Reports of criterion-referenced scores may also include degree of success or failure if such information is meritorious (Popham and Husek 1969). Statistically, NRTs offer some advantages.

Advantages of NRTs

A distinct advantage of a NRT is its compatibility with standard measurement procedures. Neel (1972, p. 219) writes that: "Standardized tests of achievement have measures of their authenticity such as validity, reliability, sampling, content and quality content audits across broad populations and consequently are truly standards of assessment." Glaser (1963, p. 520) concurs by writing: "Norm referenced measures lend themselves to work with aptitude, selection and predictions because of their compatibility with correlational analysis."

A second advantage of NRTs is that a student's score from such a test may be translated into a relative standing with his peers. Norm-referenced tests are desirable in an initial, intense evaluation of a child to determine deviation from the norm (Mann 1973). Also, Ebel (1971, p. 283) states, "In many areas of education we do pursue excellence, in many areas we are concerned with deficiency. For these purposes we need norm-referenced measures."

Some additional advantages of NRTs listed in the literature are:

1. Reasonably objective
2. Uniform
3. Standardized

4. Readily available

5. Relatively inexpensive

6. Diagnosis of general and individual needs

7. Assessment of the effectiveness of instruction.

There are also advantages to CRTs.

Advantages of CRTs

Numerous authors list one or more advantages of criterion-referenced tests. Some of the most noteworthy enumerated in the literature are:

1. Facilitates individualized instruction--considered crucial in work with learning disabled where group comparisons are of little concern (Brazziel 1972, Cheikin 1971, Hawes 1973).

2. Adaptable to all conditions (Neel 1972).

3. Congruent with teacher objectives--not a function of predetermined sets of nationally representative learning objectives and performance standards as in norm-referenced tests (Hawes 1973, Neel 1972).

4. Permits direct interpretation of progress in terms of specified behavioral objectives (Brazziel 1972).

5. Lends itself to task analysis (Neel 1972).

6. Eliminates a situation where half of American school children must always be below median as with norm-referenced tests (Brazziel 1972).

7. May provide feedback to classroom at regular intervals (Brazziel 1972, Hawes 1973, Neel 1972, Simon 1969).

8. Can service the most basic change in the instructional system (Neel 1972).

9. Eliminates pressures on teachers to "teach to the test" in order to have children make good showing (Brazziel 1972).

Disenchantment with CRTs and/or NRTs is rhapsodized in various articles.

Disadvantages of NRTs

A substantial number of authors discuss the disadvantages of NRTs. Lack of remediational guidelines is a major complaint. Block (1971) and Hawes (1973) state that NRTs provide insufficient information about a student's skill repertoire to facilitate accurate diagnosis. Such information they feel is a necessity to provide optimal remediation and avoid learning impairment. There is only description, not prescription, according to these authors.

In addition to not providing remediational guidelines, some authors feel that NRTs have a major measurement deficiency. Skagen (1971) writes that NRT scores create evaluation problems in terms of proximity to teacher or district objectives. It is conceivable that a NRT and a teacher's objectives could have little overlap.

Another dimension to the potential disadvantages of NRTs is a function of an over-zealous attempt to create variance. Popham and Husek (1969, p. 4) write: "Occasionally this overriding criterion may reduce the adequacy of the instrument for even spurious factors may be incorporated at times just to produce variance."

Another potential disadvantage of NRTs is that they can be biased against particular samples. Since most NRTs are standardized on a widespread "white middle class" population, it is felt that scores for minorities and lower socio-economic students are potentially invalid (Fleming 1971, Hawes 1973). Tyler (1971) complains that the questions on many NRTs are inappropriate for evaluating the progress of extremely slow, fast, and disadvantaged learners. His rationale is that NRTs furnish so few questions representative of their skill level that scores will not represent valid estimates of their ability. In addition, measuring student progress is difficult with such special students on NRTs because the test does not include a sufficient number of questions covering the material that they have been learning; therefore, pre-post type evaluation in these situations are often inaccurate measures of progress. The changes in scores for a student may be a chance variation since the scores are based on only a few questions which were answered from competence.

The last two disadvantages of norm-referenced measures discussed in this paper are somewhat subjective. Falcone (1969) feels that standardized test scores are necessarily categorized and labeled. This process, according to Falcone, often results in the direct or indirect divulgence of "low" performances to certain students. It is concluded that such a revelation may result in a permanent deflation of self-esteem. Also of a subjective nature, Moxley (1970, p. 53) writes: "Norm-referenced measures of evaluation have been a source of disorder in learning, but the disorders can be reduced by using measures with an

absolute or criterion measure." Moxley's syllogism seems to be based on the major premise that by proper manipulation of the environment all men can learn all things. Moxley implies such a state would be optimal. Therefore, since norm-referenced test scores indicate differences between individuals rather than differences between environments, they are contributing to educational disorder by misplacing emphasis on individual differences, instead of environmental differences. There are grievances against CRTs also.

Disadvantages of CRTs

Some writers feel CRTs represent an educational regression. Ebel (1971, p. 28) writes: "More important, the difficulties and limitations of criterion-referenced measures [e.g., defining perfect mastery] which half a century ago led to their virtual abandonment, will once again become apparent and will, in all probability, start the pendulum swinging back toward norm-referenced measures."

Hawes (1973) also states that the inability to operationally define mastery of a task is an inherent problem with CRTs. He writes that such measures may seriously mislead school personnel because they seem more precise than they really are.

Brazziel (1972) discusses four potential disadvantages of criterion-referenced measures:

1. If the tests are to be valid, materials for teaching toward specified objectives must always be available.

2. Current construct validity measurement of CRTs is not sophisticated enough.

3. Since CRT reporting systems are not standardized among districts, they must be interpreted to another district every time a student moves.

4. It is difficult to compare one district's performance with another district on the basis of CRT scores.

Ironically, Neel (1972) attributed the same disadvantagous characteristic to CRTs that was attributed to NRTs by Block (1971) and Hawes (1973), namely, CRTs provide a descriptive evaluation without a prescription. His rationale for this appears to be the fact that CRT scores cannot be compared to a "reference point," i.e., normative group scores, to evaluate need for remediation.

Additional concern about CRTs was expressed regarding test construction. Ebel (1971) hypothesizes that if objectives are to be meaningful, they must not be idiosyncratic and limited to the standards, interests, and values of just one teacher. Meetings which achieve at least verbal consensus are necessary. Ebel feels that "unproclaimed" disagreements are inherent in such a situation; therefore, the effectiveness of the instrument will be limited due to lack of general teacher enthusiasm and/or generally appropriate objectives. Other test construction problems are presented by Neel (1972):

1. Some personnel are not able to write behavioral objectives.

2. Stating educational objectives in specific operational terms may take years.

3. The item analysis necessary to generate a quality instrument may take years.

Additional disadvantages of the rationale and scope of CRTs were discussed by Ebel (1971). He writes that mastery is a matter of degree with only the simplest ability, understanding, or appreciation ever mastered completely. In conclusion, Ebel (1971, p. 283) then states: "Criteria of mastery must be imperfect and arbitrary with a consequent imperfection in a measurement instrument that utilizes such criteria." Ebel also surmises that the effective use of CRTs is limited to areas that require a high degree of skill while utilizing limited abilities. He writes that effective measurement of areas emphasizing general knowledge and understanding is probably beyond the scope of CRTs.

The last disadvantage of CRTs to be listed in this paper pertains to educational philosophy. Hawes (1973, p. 37) states: "Use of criterion-referenced tests to completely direct instruction violates the learning process and neglects vital goals." A useful index of the relative weaknesses and strengths of a test is its status among educators.

Current Status of NRTs and CRTs

After reviewing the literature on the current status of the CRT and NRT concepts among educators, it appears obvious that CRTs are beginning to preempt the popularity of NRTs. Ward (1970) writes that in the area of special education there is a discernible movement away from NRTs toward CRTs because of the diagnostic utility of the latter. Mann (1973) states that the most noticeable trend in educational measurement is the demand for improved standardized tests that permit translation

of performance data into teaching prescriptions, rather than emphasizing the normative aspects of scores. Hawes (1973, p. 38) calls CRTs the "fastest growing new technique" for measuring educational achievement. Brazziel (1972) claims that CRTs are being used as the key to evaluating students involved in the rapidly growing trend toward program individualization.

Two contemporary adaptations of criterion-referenced measures to educational needs are demonstrated in California and Florida. Under Florida's Educational Accountability Act of 1971, CRTs are currently being used to evaluate student progress in some subjects at certain grade levels. The projected use of CRTs in Florida, according to Hawes (1973, p. 36) is ". . . criterion-referenced tests at every level and in every subject." Criterion-referenced tests are considered by many educators to be the only practical way of fulfilling the student progress evaluation requirements of California's Stull Bill (see page 1 of this text). Consequently, there is an important emphasis on CRT development in California school districts.

Some authors suggest that the needs of students and educators may best be served by taking advantage of the positive qualities of both CRTs and NRTs. Hawes (1973) writes that both norm and criteria data are required for a comprehensive evaluation of students. He feels that neither norm nor criteria data alone are conclusive. Another proposal for a more eclectic approach is offered by Neel (1972) who suggests the creation of a norm-based CRT as the ultimate solution to

measuring educational progress. He provides the following reasons as rationale for such an instrument:

1. Reasonably objective
2. Uniform
3. Standardized
4. Readily available
5. Relatively inexpensive
6. Comparable to other samples of the population
7. Assist in determining whether materials and methods are effective
8. Adaptable to most conditions
9. Provides feedback to instruction through right/wrong response.
10. Reducible to basic units of instruction to the formative level of evaluation
11. Produces a pool of objectives from which teachers can choose; in addition, pool controlled by norming procedures.
12. Is congruent with objective
13. A norm-based criterion instrument is relatively easy to construct
14. Assists in deciding superior teaching methods

This increase in demand for CRTs is being heeded by educational publishers.

Published CRTs

There has been a rapid increase in the number of test publishers placing CRTs on the market (Block 1971, Brazziel 1972, Hawes 1973, Millman 1972). Many new teaching systems utilize CRT measurement for unit and course evaluation (Mann 1973).

Educational Testing Service, 1971, has developed a criterion-referenced reading test. The Psychological Corporation is currently publishing the Boehm Test of Basic Concepts which is designed in a criterion-referenced format (Hawes 1973).

The California Testing Bureau, a subsidiary of McGraw-Hill, published two of the earliest CRTs: Prescriptive Mathematics Inventory, in 1971, for grades 4 through 8, and the Prescriptive Reading Inventory, in 1972, for grades 1.5 through 6. The Prescriptive Mathematics Inventory covers 345 "specific instructional objectives" in both traditional and modern math, with one test question each. The Prescriptive Reading Inventory covers 90 objectives with three or four test questions each. In addition, for both tests, an individual lesson guide that matches unobtained objectives with text book pages is available (Hawes 1973).

Another contemporary CRT is called Diagnosis, An Instructional Aid, which is issued by Science Research Associates. It is designed for use in grades 4 through 6. This test has a one-sheet probe that enables a teacher to diagnose a pupil's progress on a continuing basis throughout the year. The teacher can utilize a combination of probe and prescription guide to plan each student's study program. In

addition the prescription guide is cross-indexed by objectives with the most commonly used mathematics texts (Hawes 1973).

Harcourt, Brace, and Janovich have published an extensive series of CRTs called the HBJ Reading Assessment System. The design of this system consists of a "location test" for original diagnosis and mini tests for continuing evaluation. The outstanding feature of the program is the "latent image" process which gives the student immediate feedback after each response. This program also includes computerized processing of scores for summary and cumulative data (Hawes 1973).

The California Testing Bureau (CTB) offers an alternative to districts that neither wish to utilize any of the commercially published tests or design their own instrument. The CTB specialists will help district personnel define their own objectives and build tests to correspond (Hawes 1973).

Objective-based Instruction

In a diagnostic teaching system, behavioral objectives are formulated. Next, a student's skill level is ascertained in relation to these behavioral objectives. After this student evaluation is completed, procedures are established that will lead to mastery of the prescribed behavioral objectives. This part of the paper discusses the construction of objectives, as well as the advantages and disadvantages of objective-based instruction, in light of the current literature.

Construction of Objectives

Basically, a behavioral objective is created by describing a goal in behavioral terms. It is also necessary to state the conditions and goal strength required to demonstrate acceptable mastery.

If objectives are to be beneficial, they must also be specific. Eustace (1969) writes that if the statement of the behavioral objective is too broad, analysis is difficult. For example, the objective "learn concept noun" would be hard to assess. However, Block (1971) warns that authors of objectives should combine specificity with prudence. He feels that objectives should be specific enough so teachers and students have sufficient direction, but that objectives should not be delineated in such great detail that they are bogged down with trivia.

Some authors feel that objectives written only in behavioral terms are too limiting. Gagne and Kneller (1972, p. 24) write: ". . . I'm inclined to argue that a complete statement of an instructional objective, designed to serve all of its communicative purposes, needs to identify both the capability learned as well as the specific performance which such a capability makes possible."

Sources of behavioral objectives are varied. They may be formulated by commercial programs, legislators, school administrators, teachers, and/or students.

Advantages

Authors suggest many advantages to objective-based instruction. Block (1971, p. 292) writes that listing of objectives doesn't stifle

creative and effective teaching as some suggest, rather ". . . it can help transform group-based instruction of any initial quality into instruction of optimal quality for each learner."

Objective-based instruction also offers the advantage of allowing more freedom for both student and teacher. Behavioral objectives measured by a CRT would allow a student to enter a course at any time and leave when mastery of certain objectives had been completed (Johnson 1971). The use of objectives also provides additional flexibility for teachers. Instructional programs that emphasize results achieved, rather than type of instruction, allow teachers to choose the technique of instruction that works best for them (Popham and Husek 1969, Trueblood 1971). There are also disadvantages to objective-based instruction.

Disadvantages

The disadvantages of an objective-based teaching system enumerated in the literature are concerned with the potential of professional stagnation. Ebel (1971) and Hawes (1973) write that general educational objectives are good, but objectives encumbered with detail on specific elements of knowledge or ability to be tested are more likely to suppress than stimulate effective teaching. Gagne and Kneller (1972, p. 24) also considered this possibility for another reason: "Behavioral objectivists can help by providing models to spur investigation. Yet if these models are adopted uncritically by the rank and file of teachers, education will decline into an unauthentic and spiritless

conditioning." Once behavioral objectives are established, a task analysis of the goal can aid instruction.

Task Analysis

Task analysis can be an important factor in the design of both diagnostic instruments and an instructional program. This section will review the basic principles and practical application of task analysis as perceived by certain contemporary educational writers.

The basic concept of task analysis is that knowledge is hierarchical. Gagne and Paradise (1961, p. 2) write, "Knowledge relevant to any given final task to be learned is conceived as a set of subordinate capabilities called learning sets."

The rationale for utilizing task analysis as an instructional tool is the positive transfer that is mediated between sequences of subordinate sets. In this regard, Gagne (1970, p. 239) writes: "A learning hierarchy, then, identifies a set of intellectual skills that are ordered in a manner indicating substantial amounts of positive transfer from those skills of lower position to connected ones of higher position."

The establishment of a valid hierarchy of learning sets is crucial for optimal diagnosis and instruction. According to Gagne and Paradise (1961, p. 16) subordinate learning sets to a performance of a task may be identified by asking: "What would the student have to be able to do in order to be able to perform this task." Gagne and Paradise also note that two or three such subordinate skills may contribute positive transfer to a higher level skill, and that a single

subordinate skill may facilitate the learning of more than one single higher level skill. It is important to note that the establishment of a learning hierarchy doesn't imply a rigid instructional format. In this regard, Gagne (1970, p. 242) writes ". . . the learning hierarchy does not picture the procedure of instruction . . . . only the internal conditions for learning, the prerequisite capabilities that will provide the positive transfer to a new learning set." Gagne elaborates that a learning hierarchy which is derived from a task analysis is analogous to a "map". Within the map there are alternate "routes" available for learning. The "route" chosen would depend on the individual learner. In other words, type of instruction, amount of guidance, etc. is not a function of task analysis.

The consequences of the deletion of training on one step of a skill hierarchy are varied. According to Gagne and Paradise (1961) and Gagne (1970), because of previously learned capabilities or innate capacities, it is possible to learn a skill without having mastered a subordinate skill. However, he also warns that in such cases the ability to generalize the higher order skill may well be limited. Having considered the principles of task analysis, practical application will be discussed next.

Okey (1970), Cunningham (1971), and Gagne (1970) write that by utilizing task analysis and the resultant learning hierarchy, a teacher will be able to diagnose the student's highest level of competency in a sequence and begin instruction there. Thus, the teacher will avoid needless repetition and frustration for the students. Maertens and

Schminke (1971) offer the following example of a sequence of skill categories that should be considered for optimal diagnosis when a teacher is evaluating a student's computational ability:

1. Use of numbers (associative learning)
2. Understanding of quantity (concept learning)
3. Synthesize concept of numeration and operation required (principle learning)
4. Arrive at answer (problem solving)

Junkola (1972, p. 33) writes that the advantage of task analysis is that it allows the instructor to maintain the student in a "state of continual readiness" for the superordinate learning set. Therefore, he concludes that making decisions about readiness is an on-going process.

Assessment

Student progress must be monitored if diagnostic teaching is to be optimally effective. Block (1971) writes that the key to diagnostic teaching is a feedback or monitoring system. In this regard, Junkola (1972, p. 38) adds: "This skill [assessment] places a teacher far ahead of those whose only explanation for failure is the child's poor memory or short attention span." The rationale for these virtues bestowed upon assessment is that feedback provides both student motivation and guidance for program development.

The best means of assessment is an instrument developed from prespecified behavioral objectives. Only with such an instrument will

validity be assured (Gagne 1970). Many of the principles of diagnostic teaching discussed in this review of literature are implemented in the study presented in the next chapter of this paper.

CHAPTER 3

THE STUDY

This chapter contains a description of a research project conducted by the author during the summer of 1972, at Ochoa Elementary School in Tucson, Arizona. The purpose for this study has been discussed in the first chapter of this paper. There are three parts to Chapter 3: materials, subjects, and procedure.

Materials

The materials used in this study were testing instruments and record keeping forms.

Instruments

The two instruments used for the study were the Heterogeneous Classroom Diagnostic Instrument (HCDI) (see Appendices A and B) and the Metropolitan Readiness Test (MRT). The HCDI was designed by the author of this paper for the purpose of this research as well as for future classroom utilization by teachers. The MRT is a standardized achievement test published by Harcourt, Brace and World, Inc.

The materials from the MRT used in the study were the mathematics and alphabet sections of form A. This author feels that it is unnecessary to describe MRT materials in this text because they are commercially available. The HCDI, however, will be discussed comprehensively.

The HCDI measures 45 different skills with 189 questions. These skills are divided into ten subtests: reading, math, concepts, grammar, labeling, demographic data, shapes, colors, cutting with scissors, and measurement, money, and time.

The questions presented in the HCDI are the result of objectives established by kindergarten, first, and second grade teachers in conjunction with the author. After the objectives were designated, task analysis was used to identify the subordinate skills necessary for mastery of each objective. The sequence of questions in a section were ordered according to a skill hierarchy where it is both possible and practical.

These questions were presented on two different forms of the HCDI-administrator's form (Appendix A) and student's form (Appendix B). There are four differences between the two forms. The student's form is typed in larger print and the questions have more generous spacing. The administrator's form includes directions and provisions for scoring. Both forms include complete copies of all the questions.

The basic rationale for the design of this test was to elicit the most accurate indication of each studen't ability, as well as provide the most useful information possible for the teaching staff. To attain valid student scores, the author incorporates into the instrument three features: individual administration, ceiling-basal scoring, and clarity of required responses. The instrument will be of benefit to teachers because scores reflect competency level of "specific"

skills and the skill hierarchy (reading and mathematics sections) facilitates remediational programs.

Record Keeping Forms

These were two types of record keeping forms used in this study. One type was used by the author to record pre- and posttest data (Appendix C). The second type was designed for the experimental classroom teacher and her aide to record this academic progress of students (Appendix D).

The classroom record keeping forms came in sets of five sheets. Each sheet was designed to record different data. The skills listed on these forms corresponded to the skills listed in the HCDI. For each skill there were seven boxes to be used to record progress and the entry date. There was a space below each box for additional notations.

These sets of forms were placed in a loose leaf notebook separated by dividers labeled with each student's name. In this way, instructional staff had easy access to each student's records.

Subjects

The population from which the sample for this study was drawn had four parameters: first, be Mexican-American; second, be in kindergarten, first, or second grade; third, have a parental request to attend summer school; fourth, meet eligibility guidelines for the federally funded Follow Through summer school program.

From this population, 20 kindergartners, 20 first graders, and 20 second graders were randomly selected. The selected students from

each grade level were then randomly divided between the experimental classroom and the control classroom.

Sixteen of the original sample of 60 students were not included in the final data analysis due to insufficient attendance at summer school or unavailability to take the posttest.

(by grade)

Experimental		Control
K	9	8
1	8	5
2	5	9
Total	22	22

(by sex)

Boys	12	11
Girls	10	11
Total	22	22

Procedure

This part of the paper describes the methodology of the study. There are three areas discussed: testor training, testing, and treatment.

Testor Training

All of the testors were trained before any tests were administered to children in the study. The first part of training consisted of comprehensive discussions between the author and each testor on the

purpose of every subtest, scoring, and directions for administering the HCDI.

The second part of training necessitated that each testor successfully administer one HCDI to a student in the same age range as the sample students. Such practice sessions were completed under the supervision of the author.

The testors were three Anglo coeds in their senior year of college at The University of Arizona. The testors were paid by the author for their time at the rate of two dollars per hour.

Pretests

The pretests of the HCDI and the MRT were administered to all students in the study on a random basis at the end of the 1971-72 school year. No student was administered both tests on the same day.

The MRTs were given to students in groups of four. The author administered all of the MRTs.

The HCDIs were administered individually by the testors, including the author. After a testor had chosen a student's name, he/she would go to the appropriate classroom to pick up the child. This early contact between student and testor facilitated rapport. All students entered the pretest sessions willingly.

All testing was done in vacant classrooms or in the library. Interruptions were minimal. All students were told they would receive a "treat" at the end of pretest sessions.

Treatment

The major difference between the instruction given to the control and experimental group was the use of a CRIS. The experimental group staff utilized the HCDI pretest score and a monitoring system to evaluate their students' needs, formulate goals, and plan remediational activities.

At the beginning of summer school, the staff for the experimental group was given a record keeping book with a set of forms for each student (see section on materials). The HCDI pretest scores, converted to percentage, had been placed next to the appropriate skills on each form. The staff used these initial scores as a base for remediational plans.

At least once every two weeks, each student's progress was assessed. Only certain skills were taught and monitored because of time constraints (academic skills were emphasized for only 45 minutes per day). The assessment procedure consisted of administering skill probes and translating the results into a percentage score, to allow for progress evaluation.

These probes were short tests designed only to evaluate a single competency. It was suggested to the instructional staff that the number of problems placed on a probe be easily convertible to percentages like ten or four; e.g., three out of four problems equal 75%.

Since none of the experimental group staff saw the HCDI until after summer school, they had no idea as to the specific content or form of the test questions. Because the probes would have to be

developed by the experimental staff without benefit of a model, the author trained them on relevant principles of probe construction.

After an experimental group student demonstrated mastery of a skill by his performance on a probe, his instructional staff formulated a new goal for him. If the skill area was reading or math, where skills were generally listed in a hierarchical sequence, the superordinate skill was the next goal.

Other than the CRIS, the experimental classroom and the control classrooms were very similar. Both classes were taught in a nonstructured manner. The classrooms were both learning-center orientated. Class objectives were unacceptable in both classes because they were perceived as being insensitive to individual differences. Also, heterogeneous grouping was practiced in both classes.

Posttests

Posttests for the HCDI and MRT were administered to all qualified experimental and control students within a week and a half after the termination of summer school. To be qualified, a student had to attend summer school for at least 80% of the days.

Students and their parents were contacted either at their homes or by mail (Appendix E) to make appointments for picking up and retesting the students. Parents were informed that the test was a follow up to a study began in the spring. The students were told that they would once again be given punch and candy upon completion of the "game". All qualified students eagerly participated.

The posttest procedures, conditions, and personnel were exactly identical with those for the pretest except that school was not in session.

CHAPTER 4

RESULTS AND DISCUSSION

The first hypothesis stated: It is feasible for a teacher or group of teachers to design and implement a CRIS which is keyed to their own educational objectives and teaching styles.

The empirical results show:

1. A CRIS was designed by the author that included a diagnostic instrument and record keeping forms.

2. The CRIS was keyed to teacher-generated objectives.

3. The CRIS was implemented into the experimental classroom instructional program.

These results indicate that the design and implementation of a CRIS is within the capability of school personnel.

The second hypothesis stated: Students in a classroom that incorporates a CRIS will progress significantly ($P < .05$) more than students in a control classroom toward mutual goals.

The results of an analysis of variance for the control and experimental group scores are presented in Table 1. The results indicate that:

1. The experimental group significantly ($P < .01$) outscored the control group on the Reading subtest.

2. There was no significant difference between the control and experimental group scores on the Math, Labeling, Demographic Data,

and Shapes subtests (the Grammar, Money and Measurement, and Color subtests were deleted because they had either an insufficient number of scores or pretest scores approximated the ceiling so closely that room for improvement was extremely limited).

TABLE 1

Analysis of Variance for
Experimental and Control Groups

Source	df	MS	F	P
Math	1	31.920	.504	.511
Reading	1	210.181	6.723	.012
Labeling	1	4.102	.114	.736
Demographic Data	1	10.920	.343	.567
Shapes	1	.920	.673	.578

These inconclusive results that appear to qualify the merit of CRIS must be considered judiciously. First, the amount of time util lized by the experimental staff for HCDI related goals was limited to only 45 minutes per day. Second, the anecdotal data obtained from the experimental instructional staff is very positive. These data include staff prepared summary of their opinion of the CRIS (Appendix F). In addition, the experimental group teacher, in what appeared to be a state of unabashed spontaneity, told this author: "I don't know how I ever got along without it [CRIS]."

APPENDIX A

HCDI - ADMINISTRATORS' EDITION

HETEROGENEOUS CLASSROOM

K-2

DIAGNOSTIC INSTRUMENT

Ochoa Elementary School

Summer, 1972

MATH

1. Counting

(Count by ones starting with one until I tell you to stop--stop at 19)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 ____

(Count by tens from ten to one hundred)
10 20 30 40 50 60 70 80 90 100 ____

2. Count Objects

(Point to the one dot on the child's test and ask how many dots there are) ____

(Point to the group of five dots and ask the child to count the dots ____

3. Locate Numbers

(In each of the following four rows, ask the child to find the numbers that are marked on the manual. Proceed a number row at a time.)

5 1 4 3 2 ____

6 0 8 9 7 ____

12 10 15 19 14 11 13 18 17 16 ____

10 30 50 70 40 90 60 20 80 100 ____

4. Identify Numerals

(Have child read row of numerals from right to left) (Circle wrong answers.)

3 2 5 1 4 ____

10 7 9 8 6 ____

15 13 11 12 14 ____

27 74 45 35 97 ____

5. Match numerals and groups of objects

(Place groups and numerals for four, eight, five, and three in front of child, and have him match them.) ____

6. Copy numerals

(Have child copy each numeral on spaces provided.)

1 2 3 4 5 6 7 8 9 10

_ _ _ _ _ _ _ _ _ _ ____

7. Recall and write numerals

(Have child write numbers that you call in space provided on his paper. Call numbers one at a time.)

4 11 0 2 3 ____

9 8 6 7 5 ____

8. Addition - Oral (sums below 10)

(Present problems orally.)

2 plus 3 = 4 plus 5 = ____

For the rest of the math section the directions are self-explanatory. ____

9. Addition - (sums below 10)

4	6
+3	+0

____

____

10. Addition - Horizontal (sums below 10)

1 + 4 = 3 + 6 = ____

____

11. Addition - (sums above 10)

8	9
+ 4	+ 7

12. Addition - Two digit without carrying

12	46
+ 14	+ 31

13. Addition - Columns of three one-digit numbers (sums below ten)

3	2
4	6
+ 1	+ 0

14. Addition - Columns of three one-digit numbers (sums above 10)

6	7
8	6
+ 4	+ 1

15. Subtraction - More and less concepts

Which number is more 6 or 2?

Which number is less 7 or 3?

16. Subtraction - One digit

7	3
- 4	- 0

17. Subtraction - Two digit no borrowing

12	34
- 8	- 21

18. Subtraction - Borrowing

17	32
- 9	- 14

READING

1. Word Attack - Letter location

(Ask child to point to the capital and lower case below as you call them off. Be sure and adhere to the following order: H, A, D, D, W, S, O, c, l, r, i, m, y)

D S W H A O

c m y l i r

2. Word Attack - Letter identification

(Have child identify the following upper and lower case letters in the order presented.)

Z N B R F Q

u a g p d m

3. Alphabet - Letter copying

(Have child copy the following upper and lower case letter in the space provided below each letter.)

B b O o F f U u

4. Alphabet - Writing from memory

(Have child write from memory the following letters in the space provided on his test.)

M R Z Q I B

a g u w f n

5. Word Attack - Phonics (consonants)

(Ask the child to give you the phonic sound of the following letters.)

T M B V R F ____

6. Word Attack - Phonics (vowels)

(Ask the child to give you the phonic sound of the following letters.)

A E I O U ____

7. Word Attack - Blends

(Ask child to give sounds of following blends.)

bl sh ch fl ____

8. Word Attack - Context clues

(Read each of the following sentences to the child. After reading each sentence ask the child to fill in the missing word. If the child doesn't know the answer, read the sentence again. Accept any answer that is reasonable in context.)

a. Henry kicked the _____ with his foot. ____

b. Ochoa is the name of my _____. ____

c. John's little furry _____ with a wagging tail followed him to school. ____

d. In the morning all of the children in class sit on the _____ . ____

e. I like Ramona because she is my _____. ____

f. Belinda has long pretty _____ that hangs down her back. ____

9. Reading Comprehension - Answering closed questions

(Have the child read each of the following questions and give the answer. Supply no help.)

a. What is the name of your school? _____ ____

b. Who is your teacher? _____ ____

c. How many eyes do you have? _____ ____

d. What is your name? _____ ____

e. How old are you? _____ ____

10. Sight Words

(Tell child to give you the name of each word that you put the marker under. Tell child that guessing is alright. Also tell child that words are difficult and you will be moving quickly so he shouldn't feel bad if he misses several.

	A ____	B ____	C ____	D ____	E ____
a.	____both	____got	____ride	____has	____these
b.	____do	____pretty	____is	____much	____make
c.	____wish	____put	____know	____see	____those
d.	____on	____pull	____not	____your	____never
e.	____best	____get	____its	____had	____for
f.	____could	____it	____kind	____must	____an
g.	____done	____does	____every	____going	____made
h.	____my	____jump	____went	____ate	____at
i.	____with	____if	____read	____yellow	____find
j.	____me	____just	____eight	____are	____fly
k.	____cold	____did	____now	____goes	____far

CONCEPTS

Materials - Five checkers, two boxes.

(Ask child the following questions to test his grasp on the underlined concepts.)

a. Which box has a pencil IN it? ____

b. Which box has a pencil ON it? ____

c. Which box has a pencil UNDER it? ____

d. Which box has a pencil BEHIND it? ____

e. Which is bigger, a bicycle or a tricycle? ____

f. Which is slower, running or walking? ____

g. Which is heavier, a ball or a chair? ____

h. Which checker is in the middle? (Out of 5 checkers in a row)

i. Which checker is first in line? ____

j. Which checker is last in line? ____

k. Which checker is second in line? ____

l. Which is more (show two checkers and five checkers)? ____

m. Which is less (show one checker and three checkers)? ____

LABELING

(Place the following items in a box. Ask the child to pull them out one by one and identify those he can. Place a check next to the ones that the child identifies correctly.)

1. chalk
2. eraser ____
3. scissors ____
4. crayon ____
5. ruler ____
6. string ____
7. glue ____
8. book ____
9. trash can (place on top) ____
10. paper ____
11. paper towel ____
12. saw ____
13. hammer ____
14. jump rope ____
15. Kleenex ____
16. scotch tape ____
17. paper punch ____
18. spoon ____
19. pan ____
20. glass ____
21. toothbrush ____
22. thumb tack ____
23. thumb tack ____
24. knife ____
25. paint brush ____
26. cup ____
27. paper clip ____
28. record ____

DEMOGRAPHIC DATA

(Have each child fill in the following information on their test.)

First Name _____ ____

Last Name _____ ____

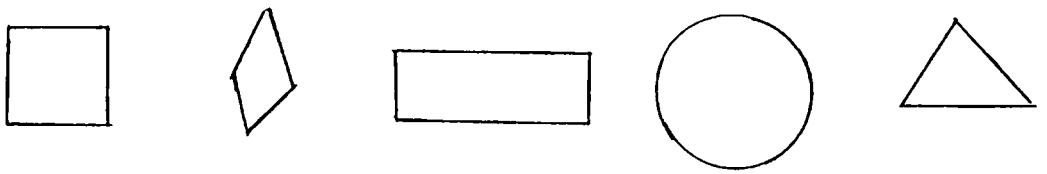
Telephone Number _____ ____

Address _____ ____

SHAPES

1. Identification

(Ask child to point to whatever shape the question calls for.)



a. Which shape is a circle? ____

b. Which shape is a square? ____

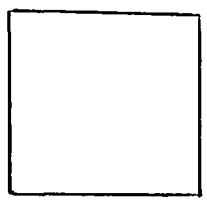
c. Which shape is a rectangle? ____

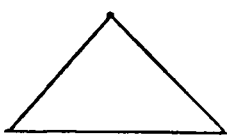
d. Which shape is a triangle? ____

e. Which shape is a diamond? ____

2. Copying

(Ask child to copy the square and the triangle.)





COLORS

Materials - red, black, yellow, blue, and green crayons.

(Ask child the following questions)

a. Which is the red crayon? ____

b. Which is the green crayon? ____

c. Which is the black crayon? ____

d. Which is the blue crayon? ____

d. Which is the yellow crayon? ____

CUTTING WITH SCISSORS

(Have child cut along both straight line and outline of circle on his test.)

Straight line ____

Circle ____

DICTIONARY

(Have child find word dog in dictionary.) ____

GRAMMAR

(Have child place any periods, commas or question marks that are necessary in the following sentences. It is permissible to inform the child of the names of any words that he doesn't know.)

a. I am a boy ____

b. I like horses dogs cats and rabbits ____

c. Do you like ice cream ____

ALPHABETIZING

(Put the following words in alphabetical order, using numbers 1-5 to show what the order is.)

_____ you	_____
_____ me	_____
_____ be	_____
_____ and	_____
_____ play	_____

MEASUREMENT, MONEY, AND TIME

1. Length - Inches

   (Measure pencil or pen that is six inches long with ruler.) _____

2. Length - Feet

   (Measure top of desk in feet) _____

3. Volume - Cups

   (Ask child to draw two cups of water from a full container and place them into an empty bowl.) _____

4. Time - Daily schedule

   (Have child give times of following:) *Different schedule for K-2

   Beginning of school _____

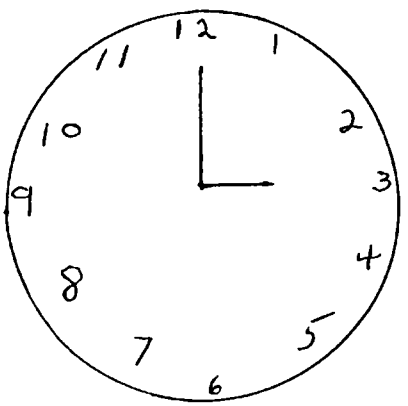
   Lunch time _____

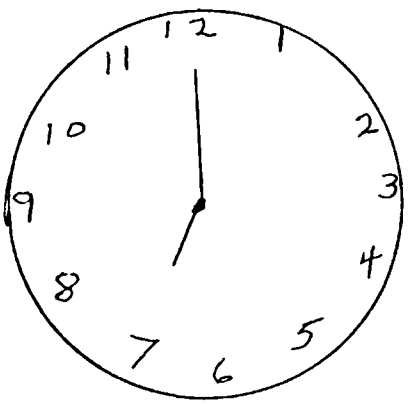
   School is over _____

5. Time - Clocks

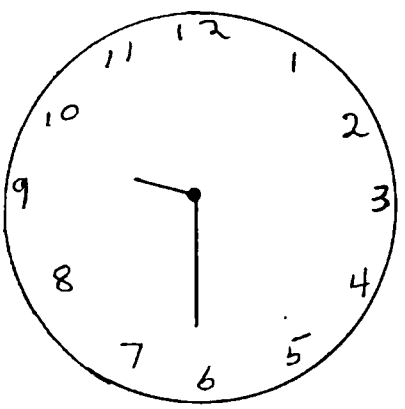
(Have child read time as shown by the clocks)

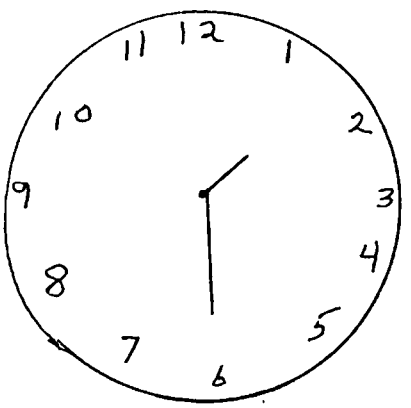
HOURS





HALF HOURS





6. Money - Identification

(Place penny, nickle, dime, and quarter in front of child and have him locate each upon request.)

Penny ____

Nickle ____

Dime ____

Quarter ____

7. Money - Relative value

(Ask child the following questions:)

a. Which is more, a penny or a dime? ____

b. Which is more, a nickle or a dime? ____

c. Which is more, a dime or a quarter? ____

d. How many pennies in a nickle? ____

e. How many pennies in a dime? ____

f. How many pennies in a quarter? ____

g. How many nickles in a dime? ____

h. How many nickles in a quarter? ____

APPENDIX B

HCDI - STUDENT COPY

HETEROGENEOUS CLASSROOM

K-2

DIAGNOSTIC INSTRUMENT

Ochoa Elementary School

Summer, 1972

Compiled by:

Huntley Hoffman
Barbara Tallis
K-2 Teaching Staff

STUDENT COPY

STUDENT NAME _____

GRADE _____

GROUP _____

MATH

1. COUNTING

2. COUNT OBJECTS

o

o o o o o

3. LOCATE NUMBERS

5 1 4 3 2

6 0 8 9 7

12 10 15 19 14 11 13 18 17 16

10 30 50 40 90 60 20 80 100

4. IDENTIFY NUMERALS

3 2 5 1 4

10 7 9 8 6

15 13 11 12 14

27 74 45 35 97

5. MATCH NUMERALS AND GROUPS OF OBJECTS

6. COPY NUMERALS

1 2 3 4 5 6 7 8 9 10

_ _ _ _ _ _ _ _ _ _

7. RECALL AND WRITE NUMERALS

____ ____ ____ ____ ____ ____

____ ____ ____ ____ ____ ____

8. ADDITION - ORAL (sums below 10)

2 plus 3 =

4 plus 5 =

9. ADDITION - (sums below 10)

4	6
+ 3	+ 0

10. ADDITION - HORIZONTAL (sums below 10)

1 + 4 =

3 + 6 =

11. ADDITION - (sums above 10)

8	9
+ 4	+ 7

12. ADDITION - TWO DIGITS WITHOUT CARRYING

12	46
+ 14	+ 31

13. ADDITION - COLUMNS OF THREE ONE-DIGIT NUMBERS (sums below ten)

3	2
4	6
+ 1	+ 0

14. ADDITION - COLUMNS OF THREE ONE-DIGIT NUMBERS (sums above 10)

6	7
8	6
+ 4	+ 1

15. SUBTRACTION - MORE AND LESS CONCEPT

6	2
7	3

16. SUBTRACTION - ONE DIGIT

7	3
- 4	- 0

17. SUBTRACTION - TWO DIGIT NO BORROWING

12	34
- 8	- 21

18. SUBTRACTION - BORROWING

17	32
- 9	- 14

READING

1. WORD ATTACK - LETTER LOCATION

D S W H A O

c m y l i r

2. WORD ATTACK - LETTER IDENTIFICATION

Z N B R F Q

u a g p d m

3. ALPHABET - LETTER COPYING

B b O o F f U u

____ ____ ____ ____ ____ ____ ____ ____

4. ALPHABET - WRITING FROM MEMORY

____ ____ ____ ____ ____ ____

____ ____ ____ ____ ____ ____

5. WORD ATTACK - PHONICS (consonants)

T M B V R F

6. WORD ATTACK - PHONICS (vowels)

A E I O U

7. WORD ATTACK - BLENDS

bl sh ch fl

8. WORD ATTACK - CONTEXT CLUES

a. Henry kicked the _____ with his foot.

b. Ochoa is the name of my _____.

c. John's little furry _____ with a wagging tail followed him to school.

d. In the morning all of the children in class sit on the _____.

e. I like Ramona because she is my _____.

f. Belinda has long pretty _____ that hangs down her back.

9. READING COMPREHENSION - ANSWERING CLOSED QUESTIONS.

a. What is the name of your school?

b. Who is your teacher?

c. How many eyes do you have?

d. What is your name?

e. How old are you?

10. SIGHT WORDS

both	got	ride	has	these
do	pretty	is	much	make
wish	put	know	see	those
on	pull	not	your	never
best	get	its	had	for
could	it	kind	must	an
done	does	every	going	made
my	jump	went	ate	at
with	if	read	yellow	find
me	just	eight	are	fly
cold	did	new	goes	for

CONCEPTS

LABELING

DEMOGRAPHIC DATA

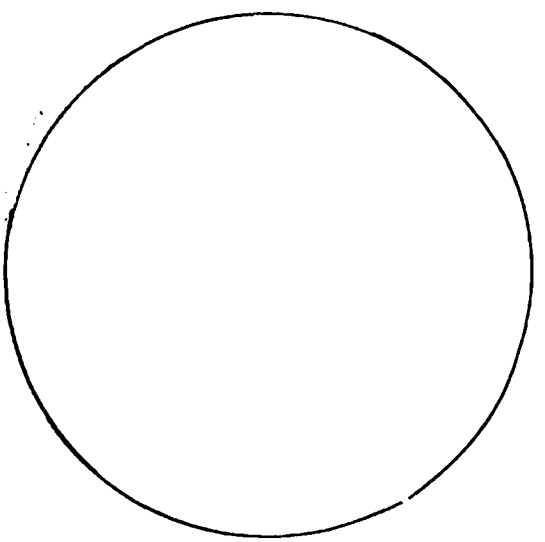
First Name _____

Last Name _____

Telephone Number _____

Address _____

CUTTING WITH SCISSORS



DICTIONARY

Dog

GRAMMAR

a. I am a boy

b. I like horses dogs cats and rabbits

c. Do you like ice cream

ALPHABETIZING

_____ you

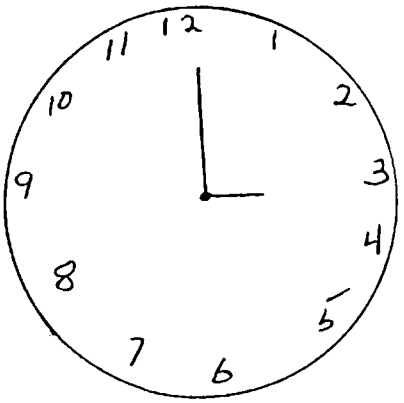
_____ me

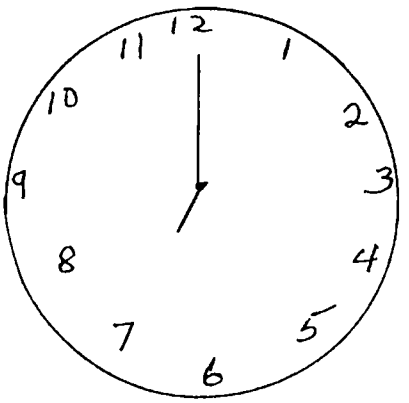
_____ be

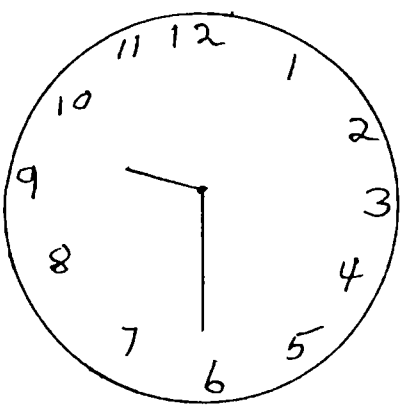
_____ and

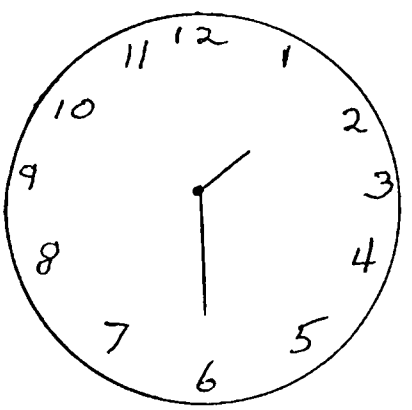
_____ play

TIME - CLOCKS


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11

APPENDIX C

PRE-/POSTTEST DATA KEEPING FORM

MATH ____|____

1. ____|____
2. ____|____
3. ____|____
4. ____|____
5. ____|____
6. ____|____
7. ____|____
8. ____|____
9. ____|____
10. ____|____
11. ____|____
12. ____|____
13. ____|____
14. ____|____
15. ____|____
16. ____|____
17. ____|____
18. ____|____

READING ____|____

1. ____|____
2. ____|____
3. ____|____
4. ____|____
5. ____|____
6. ____|____
7. ____|____
8. ____|____

CONCEPTS

____|____

LABELING

____|____

DEMOGRAPHIC DATA

____|____

SHAPES

____|____ ____|____

COLORS

____|____

CUTTING WITH SCISSORS

____|____

DICTIONARY

____|____

GRAMMAR

____|____

ALPHABETIZING

____|____

MEASUREMENT, MONEY, TIME ____|____

1. ____|____
2. ____|____
3. ____|____
4. ____|____
5. ____|____
6. ____|____
7. ____|____
8. ____|____

Name _____

Group _____

Test _____

APPENDIX D

STUDENT PROGRESS FORM

DIPC-Individual Record
Form No. 1

Name _____

Subjects Skills	Date*	Score**	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score
MATH																
Counting																
Special attention (SA)																
Counting objects																
SA																
Locating numerals																
SA																
Identifying numerals																
SA																
Match numerals and groups of objects																
SA																
Copying numerals																
SA																
Recalling and writing numerals																
SA																
Addition-oral presentation (sums below 10)																
SA																
Addition-vertical (sums below 10)																
SA																

* Confine to top box for each skill

** Scores should be recorded in terms of percent. To facilitate computing percentages, sample sizes should be in multiples of five when possible.

DIPC-Individual Record
Form No. 2

Name _____

	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score
MATH (Continued)														
Addition-horizontal (sums below 10)														
SA														
Addition-vertical (sums above 10)														
SA														
Addition-vertical two digit without carrying														
SA														
Addition-columns of three one-digit numbers (sums below 10)														
Special attention (SA)														
Addition-columns of three one-digit numbers (sums above 10)														
SA														
Subtraction-more and less concepts														
SA														
Subtraction-one digit														
SA														
Subtraction-two digit without borrowing														
SA														
Subtraction-two digit with borrowing														
SA														

DIPC-Individual Record
Form No. 3

Name _____

Subjects	Skills	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score
READING															
	Word attack-letter location														
	SA														
	Word attack-letter identification														
	SA														
	Alphabet-letter copying														
	SA														
	Alphabet-writing letters from memory														
	SA														
	Word attack-phonics (consonants)														
	SA														
	Word attack-phonics (vowels)														
	SA														
	Word attack-blends														
	Special attention (SA)														
	Word attack-context clues														
	SA														
	Reading comprehension-answering closed questions														
	SA														
	Sight words														
	SA														

DIPC-Individual Record
Form No. 4

Name _____

Subject Skills	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score
CONCEPTS														
Conceptual understanding														
SA														
LABELING														
Naming classroom objects														
SA														
DEMOGRAPHIC DATA (writing ability)														
First name														
SA														
Last name														
SA														
Telephone number														
SA														
Address														
SA														
SHAPES														
Identification														
Special attention (SA)														

DIPC-Individual Record
Form No. 5

Name _____

Subject Skills	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score
SHAPES (Continued)														
Copying														
SA														
COLOR														
Location														
SA														
CUTTING WITH SCISSORS														
Cutting on line														
SA														
DICTIONARY														
Locating words														
SA														
GRAMMAR														
Period placement														
SA														
Comma placement														
SA														

DIPC-Individual Record
Form No. 6

Name _____

Subject	Skills	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score
GRAMMAR (Continued)															
	Question mark placement														
	SA														
ALPHABETIZING															
	Placing words in alphabetical order														
	Special attention (SA)														
MEASUREMENT															
	Length in inches														
	SA														
	Length in feet														
	SA														
	Volume by cups														
	SA														
TIME															
	Important times of daily schedule														
	SA														
	Hour identification														
	SA														

DIPC-Individual Record
Form No. 7

Name _____

Subject	Skills	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score	Date	Score
TIME (Continued)															
	Half-hour identification														
	SA														
MONEY															
	Location														
	SA														
	Relative value														
	SA														

APPENDIX E

PARENTAL PERMISSION REQUEST

July 25, 1972

Dear Parents:

We are glad that your child was able to attend our Follow Through summer program this year. We hope that we will be able to continue offering such a summer program at Ochoa for our students.

To assist Ochoa to remain eligible for the Follow Through money necessary for summer school, we want to show how much our children learned during the summer program. In order to do this, we need each child who went through summer school in Barbara Tallis' classroom to come to Ochoa for about an hour as soon as summer school is over. Please indicate below the date (any day after July 30) and time (9:30 A.M.-7:00 P.M.) that your child can come to the Follow Through office at Ochoa School. We need your cooperation in getting your child to school for just one hour. If you would like us to pick your child up, please indicate below. Please have your child bring this paper back to school on Thursday, July 27, 1972. For more information, call 791-6833.

Sincerely,

Mrs. Rosalie Aros, Follow Through Coordinator

Mr. Huntley Hoffman, Psychologist

Estimados Padres:

Tuvimos mucho gusto que su nino (nina) pudo atender el programa de verano de Follow Through. Ojala que puedamos siguir ofreciendo el programa de verano para los ninos en la Escuela Ochoa.

Para poder ayudar a la Escuela Ochoa que siga siendo elegible para recibir fondos necesarios para el programa de verano, queremos ensenar cuanto aprendieron los ninos en dicho programa. Para poder hacer esto, necesitamos que cada nino que vino al programa de verano en la clase de la Sr. Barbara Tallis, venga a la Escuela Ochoa por una hora despues de que el programa haya terminado. Favor de indicar abajo la fecha (cualquier dia despues del dia 30 de julio) y la hora (de 9:30 A.M.-7:00 P.M.) que su nino (nina) pueda venir a la oficina de Follow Through en la Escuela Ochoa. Necesitamos su cooperacion para que su nino venga a la escuela por una hora. Si quiere que recogjamos a su nino, favor de indicar abajo. Favor de devolver este papel para el jueves, 27 de julio. Para mas informacion, llame al 791-6833.

Sinceramente,

Mrs. Rosalie Aros, Follow Through Coordinator

Mr. Huntley Hoffman, Psychologist

CHILD'S NAME (NOMBRE DEL NINO)_____

DATE (FECHA)_____ TIME (HORA)_____

Transportation for child needed. (Necesito transportacion para el nino.)

Transportation for child not needed. (No necesito transportacion para el nino.)

APPENDIX F

STAFF EVALUATION OF CRIS

CRIS evaluation by experimental group instructional staff:

1. Good indicator of where the child is--at what level he is working and what experiences are lacking.

2. Helps teacher and teacher aide to plan experiences that meet more individualized needs and to move the child at his own rate.

3. Concise, effective way of keeping records. Allows the teacher and teacher aide to be specific in evaluating skills. It was helpful and time saving to have all the child's records in one place and divided into subject areas.

4. There is no standard expectation for the classroom. Each child is allowed to develop and expand to his fullest potential.

5. Need space to make narrative notations of a child's growth in the other 3 goal areas--language, motivation, and intellectual base.

   ex: Belia took pride in understanding a concept.

SELECTED BIBLIOGRAPHY

Block, James H. Criterion-referenced measurement. Potential School Review, 1971, 79 (2), 289-297.

Bloom, Benjamin S. Learning for mastery. Evaluation Comment, 1968, 1, 24-32.

Bloom, Benjamin S., Hastings, Thomas J., and Madans, George F. Handbook of Formative and Summarative Evaluation of Student Learning. New York: McGraw Hill, 1971.

Bond, Guy L. Diagnostic teaching in the classroom. Reading Diagnostic and Evaluation, 1972, 13 (4), 126-138.

Brazziel, William F. Criterion-referenced tests. Today's Education, 1972, 61, 52-53.

Cheikin, Martin. Standardized tests revisited. The School Counselor, 1971, 36, 71-72.

Christoplos, Florence. Prescriptive teaching--a cognitive approach to social and interpersonal learning. Teaching Exceptional Children, 1970, 2, 158-162.

Cronbach, Lee. How can instruction be adapted to individual differences? In Robert Gagne (Ed.), Learning and Individual Differences. Columbus: Charles E. Menel Books, Inc., 1967. Pp. 20-31.

Cummings, Stanley L. Testing as a learning tool and involvement strategem. Science Teacher, 1972, 39, 49-50.

Cunningham, Donald J. Task analysis and part versus whole learning methods. Audio Visual Communication Review, 1971, 19, 365-398.

Ebel, Robert L. Criterion-referenced measurement. School Review, 1971, 71 (2).

Eustace, Barbara W. Learning a complex concept at differing hierarchical levels. Journal of Educational Psychology, 1969, 60 (b), 449-452.

Falcone, John F. The diagnostic classroom. Catholic Educational Review, 1969, 65 (2), 139-144.

Feely, Mary S., Hoyle, Virginia B., and Stanffer, Hilda. Beyond the standardized test: Objective-based evaluation in reading. Journal of Reading, 1972, 15, 616-617.

Fleming, Margaret. Standardized tests revisited. School Counselor, 1971, 19, 71-72.

Gagne, Robert M. The Conditions of Learning. New York: Holt, Rinehart, and Winston, Inc., 1970.

Gagne, Robert M., and Kneller, George T. Behavioral objectives? Yes or no? Educational Digest, 1972, 37, 24-27.

Gagne, Robert M., and Paradise, Noel E. Abilities and learning sets in knowledge acquisition. Psychological Monographs, 1961, 75 (Whole No. 518), 1-23.

Glaser, Robert. Instructional technology and the measurement of learning outcomes. American Psychologist, 1963, 18, 519-21.

Glaser, Robert. Adopting the elementary school curriculum to individual performance. Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton: Educational Testing Service, 1968.

Glaser, Robert. Instructional technology and measurement of L outcomes: Some questions. In W. James (Ed.), Criterion Reference Measurement: An Introduction. New Jersey: Educational Technology Publications, 1971. Pp. 5-14.

Harris, Chester W. An interpretation of Livingstons' referenced tests. Journal of Educational Measurement, 1972, 9, 27-29.

Hawes, Gene R. Criterion-referenced testing. Nations Schools, 1973, 91 (2), 35-41.

Johnson, Rita B. Objectives based accountability procedures for classroom use. Educational Technology, 1971, 11, 49-51.

Junkola, John. Task analysis and instructional alternatives. Academic Therapy, 1972, 8 (1), 33-39.

Larsen, Edwin P. Why test? Thrust for Education Leadership, 1972, 1 (6), 21-26.

Livingston, Samuel A. Criterion-referenced application of classical test theory. Journal of Educational Measurement, 1972, 9, 13-26.

Maertens, Norbert, and Schminke, Clarence. Teaching for what? The Arithmetic Teacher, 1971, 18, 449-456.

Mann, Lester, and Proger, Barton B. Criterion-referenced measurement: The world of gray versus black and white. Journal of Learning Disabilities, 1973, 6, 72-84.

Massad, Carolyn. Interpreting and using test norms. The Reading Teacher, 1972, 26 (3), 284-285.

McCarthy, William G. Diagnosis, prescription, connection and evaluation in the teaching of reading. Contemporary Education, 1971, 42, 233-235.

Miller, Jack. Individualizing instruction through diagnosis and evaluation. Childhood Education, 1970, 36 (8), 417-421.

Millman, Jason. Criterion-referenced measurement: An alternative. The Reading Teacher, 1972, 26, 278-81.

Moxley, Roy A. A source of disorder in the school and a way to reduce it: Two kinds of tests. Teacher and Technology Supplement, 1970, 23-26.

Neel, Thomas. Classroom performance standards. Thrust for Education Leadership, 1972, 2 (1), 17-20.

Okey, James R. Diagnosing learning difficulties. Science Teacher, 1970, 37, 59-61.

Popham, W. James. Performance tests of teaching proficiency, rationale, development, and validation. American Educational Research Journal, VIII (January), 1971, 105-117.

Popham, W. James, and Husek, T. R. Implications of criterion-referenced measurement. Journal of Educational Measurement, 1969, 6, 1-9.

Prillaman, Douglas, and Prouty, Robert W. Diagnostic teaching: A modest proposal. The Elementary School Journal, 1970, 2, 265-270.

Simon, George B. Comments on implications of criterion-referenced measurement. Journal of Educational Measurement, 1969, 6 (4), 259-260.

Skagen, R. W. System for objective-based evaluation--reading. Evaluation Comment, Center for the Study of Evaluation, 1971, 3 (1), 6-11.

Smith, Carl. Editorial. Viewpoints, 1972, 48 (1), 1-3.

Trueblood, Cecil R. A model for using diagnosis in individualizing mathematics instruction in the elementary school classroom. The Arithmetic Teacher, 1971, 17, 505-511.

Tyler, Ralph W. Testing for accountability. Education Digest, 1971, 36, 12-14.

Ward, J. On the concept of criterion-referenced measurement. British Journal of Educational Psychology, 1970, 40, 314-23.