BALL STATE UNIVERSITY

EFFECTIVE EVALUATION IN MATHEMATICS

ID 499
Senior Honors Project
Under the Supervision of Dr. Alice Robold
Department of Mathematics

March 26, 1969

Susan B. Hammersley

I recommend this thesis for acceptance by the Honors Program of Ball State University for graduation with honors.

Alice L. Robold
Thesis Adviser

Department of Mathematics

April 21, 1969
(date)

Preface

Feeling that there is a lack in the effective evaluating of the really important outcomes in mathematics learning and being especially concerned with my personal inadequacy in this field, I have undertaken to deepen my understanding and gain insight into the problem of evaluation. I have attempted to gain information and stimulation of thought through the reading of various sources, in evaluation in general, mathematics evaluation, and instruction in mathematics, and through direct experience with children.

Many experiences and people have contributed at least indirectly to the writing of my paper. I would acknowledge those who have most directly made my writing possible.

To Dr. Alice Robold I am most grateful for the inspiration and awareness imparted to me during a mathematics methods and materials course, which she taught, and for her guidance and encouragement in advising me in this present work.

My appreciation and gratitude are extended to the Frankfort, Indiana School System for giving me the full utilization of a classroom in the Lincoln Elementary School and all the facilities, including the library, gymnasium, film equipment, and other visual and manipulative aids. I give special thanks to Mr. Harry Moore, the principle of Lincoln School, who was helpful in my selection of children and in making all the school facilities and materials available for our use.

I also wish to thank four children, Dana, Dondra, Kenneth, and William for making my project not only possible but most enlightening.

The large body of this paper pertaining to readings in evaluation was made possible through the use of the libraries of Ball State University and Butler University.

Table of Contents

Introduction

Contemplating an improvement of some kind, whether it be in an industrial machine, an educational program, or in one's own personality, a person makes a decision by drawing from the experience and thoughts of others as well as those of his own. Such has been my procedure as I recognized the need for improvement in mathematics evaluation.

Beginning with many ideas gathered from various educatiors with whom I had come in contact, I broadened my knowledge by reading many sources concerning evaluation in education and particularly in mathematics. My direct experience was furthered by actually working with children. Throughout the project, direct and vicarious experience led me to do much thinking and opened new paths of thinking, which in turn directed me to further reading and experimentation.

It has not been my intention to derive a new theory in evaluation or an earthshaking discovery in an evaluation technique. I have sought to bring attention to the thoughtful analysis of evaluation and the advocation of techniques of evaluation of many previous years to point out that we are behind in our practices of evaluation. However admirable our intentions in evaluation, we often do not follow through in our actual practice of evaluation.

This paper contains four main sections. The two chapters on related readings are documentation of various thoughts concerning what evaluation should be and do and suggesting means of evaluating more effectively.

The third section describes the purpose and method of my project with children. Included are samples of anecdotal records to help demonstrate how a certain technique may be used, the responses it may evoke, and how understanding of a child may be gained through such a procedure. I have described several techniques, none of which may be new, to suggest that these have been used with some success in evaluation and to reveal the simple materials required.

My concluding statements are a result of my reading and experimenting, but are merely the beginning of my thinking and action to continuously strive toward effective evaluation in mathematics.

RELATED READINGS

IN THE AREA OF

EDUCATIONAL EVALUATION

RELATED READINGS IN THE AREA OF

EDUCATIONAL EVALUATION

Changes in the Concept of Evaluation

In reading in the area of evaluation in education we find that change has come at least in theory, if not in actual practice, in educators' concept of education which has called for changes in instruction and evaluation.

Togerson and Adams contribute the emergence of the modern concepts of evaluation to 3 main factors; a) the testing movement, from the beginning of the century through the 1920's, which helped arouse the profession to the extent and significance of individual differences; b) the inadequacy of early tests and the misuse of test results, leading to certain undesirable outcomes; and c) the curriculum change and the child-study approach leading to needed modifications in evaluation. They believe that the tests developed in the 30's to measure understandings, attitudes, and appreciations as well as skills and information made inevitable a broader concept of evaluation. (1)

Shane and McSwain, in considering the origins of the concepts of evaluation as a "phase of man's effort to apply intelligence to the rational solution of his problems," speak of the contributions of philosophy, with regard to Butler's writing of pragmatism and Dewey's method of inquiry, and of science, with its search-discovery method in many ways anticipating current evaluation procedures in which children share. They cite a) the scientific movement, providing many tools and measuring techniques of use in evaluative appraisal, and b) child development research as the educational trends influencing evaluation. Child development research emphasized the need to reexamine values as knowledge accumulated, and scrutiny of these values led to dissatisfaction with certain beliefs and practices of education. (2)

Wrightstone et al. distinguish the trends in evaluation in recent years as 1.) the concern of the modern teachers and supervisor with important functional learning outcomes, 2.) increased emphasis on the measurement of understanding and its interpretation rather than on isolated information, skills, and abilities, 3.) increased use of informal, or teacher-made, test exercises for instructional purposes to supplement formal or standardized tests, 4.) development of factor analysis of mental abilities, 5.) the development of techniques for measuring the role of individuals, as well as of small groups, in the study of group relations, and 6.) the increased attention given to the development and refinement of unstructured, or projective, tests of personality. (3)

Ahman and Glock, and Baron and Bernard in addition to the authors previously cited call attention to the development of a variety of techniques for evaluation corresponding to the emphasis on the personal and social development of the child, as well as on academic achievement, and stress the use of informal techniques. (4,5)

Modern evaluation differs from older forms of appraisal, according to Wrightstone et al., in 1.) attempting to measure a comprehensive range of objectives of the modern school curriculum rather than subject matter achievement only, 2.) using a variety of techniques of appraisal, and 3.) integrating and interpretting various indices of behavior into an inclusive portrait of an individual or educational situation. (6)

The Nature of Evaluation

The term evaluation has gained wide currency, especially since 1940, but despite its frequent use, evaluation has failed to acquire a clear cut meaning.

It is therefore helpful to consider various definitions, characteristics, and purposes of evaluation suggested by several authors in an attempt to gain a fuller understanding of the compass of evaluation.

Attempting to define and characterize evaluation as it should be today we might first distinguish between evaluation and measurement. Ahman et al. state that evaluation is a more inclusive process; measurement being that part of evaluation which provides information upon which the evaluation is based. Thus "educational measurement attempts to obtain a quantified representation of the degree to which a trait is possed while educational evaluation is a process in which a teacher commonly uses information derived from many sources to arrive at a value judgment. This evaluation may or may not be based upon measurement data." The major purpose of evaluation, say the writers, is to help the classroom teacher ascertain the degree to which educational objectives have been achieved by pupils and help him obtain better knowledge of pupils as individuals. (7)

Wrightstone et al. cite Monroe in distinguishing between measurement and evaluation by indicating that in measurement the emphasis is on single aspects of subject-matter achievement or specific skills and abilities, whereas in evaluation, emphasis is on broad personality changes and major objectives of the educational program. They further elaborate that evaluation involves 1.) identification and formulation of a comprehensive range of major objectives of a curriculum, 2.) their definition in terms of pupil behavior to be realized, and 3.) the selection or construction of valid, reliable and practicable instruments for appraising major objectives of the educational process or characteristics of growth and development. Measurement, they state, through the use of achievement tests, yields a measure of pupil attainment in subject-matter areas, especially the acquisition of skills and information. (8)

When measurement data for a child or class are viewed as the basis for judging the goodness or worth-whileness or adequacy of a child's achievement or adjustment, according to Torgerson and Adams, evaluation is taking place. The quality of a child's or groups achievement or adjustment is determined by relating measurement data 1.) to the goals instruction as established by the society, and 2.) to the needs of the child as determined by his own maturity level, his own drives and purposes, and the environmental pressures which affect him. (9)

Shane and McSwain help the reader gain insight into the nature of evaluation by providing definitions of several different educators.

> Evaluation ...implies a process by which the values of an enterprise are ascertained. (Smith and Tyler)

> A more adequate concept of evaluation includes all the activities whereby an individual or group determines how well agreed-upon purposes have been achieved. (Alexander and Saylor)

Evaluation is a process of gathering and interpretting evidence on changes in behavior of students as they process through school. (Quillen and Hanna)

The process of determining the extent to which values are achieved, purposes carried out and goals reached is evaluation. The term is derived from the word 'value', since appraisal of the effectiveness of an educational experience should be based upon a consistant , sound, democratic system of values. Evaluation includes the analysis of the purposes themselves and the consideration of techniques by which goals are attained, as well as the degree of achievement... (Burr, Harding, and Jacobs) (10)

Shane and McSwain, being particularly interested in evaluation's relation to the elementary school curriculum, distinguish between externalized and internalized evaluation. Externalized evaluation is the process whereby a group of citizens and teachers may gauge the success with which a school is achieving its purposes. Internalized evaluation is a process within us, as a result of which we make our interpretations of the environment about us, and, accordingly, direct our behavior. This is important for the evaluator to remember. The ultimate test of the external evaluation process is the degree to which it motivates the child to create within himself those personal goals that are similar to those established by the school. The goal of externalized evaluation can be realized only if the internal well springs of a childs behavior are changed by the evaluation process. (11)

Baron and Bernard list elements of a good appraisal practice as follows:

1. Pupil evaluation is a means of communication between school, child and home. As such, it must be meaningful to all concerned.

2. The purpose of evaluation is to promote optimum growth. An indication of status is not enough.

3. Evaluation should indicate what steps should be taken next.

4. Appraisal should be in terms of individual accomplishment, not in terms of interpersonal comparisons.

5. Evaluation should be in terms of stated objectives of education for the school level concerned.

6. Evaluation should be a continuing process, not an end in itself.

7. Objective data are necessary, but these data always relate to the living, dynamic person.

8. Alterations of evaluation procedures involve the entire philosophy of the school and must therefore be a matter for serious study. (12)

Rothney re-emphasizes that evaluation, to be effective must be a continuous, cooperative, and cumulative procedure. He reminds us that if the purpose of our instruction is to produce certain changes in pupils, we should find out whether those changes are taking place. We should decide what important changes we are trying to produce in children, state how changes may be expected to be shown in pupils' behavior and measure the changes that take place. Then we must study and interpret observed changes so reports may be meaningful to pupils, parents, and anyone concerned. (13)

The steps in evaluation are recognized by Shane as:

1. The isolation and description of the problem (s) to be studied

2. Clarification of the values bearing on the problem (s)

3. Development of criteria for the study of the problem (s)

4. Expression of criteria in terms of behavior sought in the child

5. Establishment of a situation in which the child's behavior can be studied as the school seeks to modify it to conform with desired outcomes

6. Employment of instruments to gather behavioral data to be studied in order to decide whether or not a child's behavior is being changed significantly

7. Analysis of behavioral changes, if significant, to determine whether or not these changes are compatible with values sought by the school

8. The implementation of decisions reached in view of the findings made by the school staff (14)

Shane lists characteristics of evaluation as he states:

1. Evaluation is an inclusive process; all available means of gathering pertinent data bearing on child behavior may be used.

2. Each situation in which evaluation occurs is unique. The teacher must begin with the background peculiar to the group in order to guide changes in behavior and decide ways in which values are to be sought.

3. Evaluation is a continuous process.

4. Evaluation is a useful means to the study of intangibles, as well as overt behavior.

5. Evaluation is concerned with total situations; related qualities and elements are studied together.

6. Evaluation is a group endeavor. All involved in examining and improving behavior participate in evaluation and the evaluation operation is in the context of human values, which requires group action.

7. Evaluation is not neutral, but biased by the evaluator's values that determine the nature of the behavioral goals or citizenship outcomes.

8. Sought values determine the outcomes in evaluation as well as the nature of the procedure used.

9. Evaluation is a contributing factor in the rational solution of a persistent educational problem.

10. Evaluation clarifies goals for education. Teachers and children cannot consciously appraise their work without developing a greater awareness of reasons for actions they take; objectives become better understood and more acceptable. (15)

The Evaluation Program

Ahman and Glock, Baron and Bernard, Torgerson and Adams, and Wrightstone et al. devote at least a chapter in their books to the problem of planning and administering the evaluation program. Shane and McSwain and Rothney, as well as the above, seem to agree on the most important considerations to be taken in planning the evaluation program which may be summed up in the steps Wrightstone et al. give.

1. Formulation of the major objectives of the curriculum. These become guideposts in curricular development and evaluations.

2. Definition and clarification of the major objectives – should outline skills, abilities, understandings, attitudes, and interest to be achieved.

3. Selection of available tests or measures and determination of the appropriateness of the selections for each of the major objectives.

4. Construction of tests, scales, or techniques needed to evaluate objectives for which no standard measuring devices are available.

5. Application of various formal and informal tests and techniques in order to make judgments about individual and group growth and development in each of the major objectives. (16)

Wrightstone and his co-writers pose several crucial questions which should help appraise the adequacy of a program of evaluation.

1. Is the design of the evaluation program comprehensive so that it includes not only abilities, skills, and understandings, but also the less tangible objectives of learning and instruction?

2. Are changes in the behavior of individual the basis for evaluating his growth and development, since total behavior of the individual should be the concern of the teacher and supervisor in every situation?

3. Are results of evaluation organized into meaningful interpretations so that a portrait of an individuals growth and development and the interrelationship of such growth become evident?

4. Is the evaluation program continuous and interrelated with curriculum development? (17)

Torgerson and Adams quote Wrightstone's characteristics of a good evaluation program stated in <u>Appraisal of New Elementary School Practices</u>.

1....evaluation must...be compatible with purposes...

2....a program of evaluation be comprehensive. It should not be limited to a few isolated goals, or objectives, but should include all major objectives of instruction.

3....for a variety of major objectives of instruction no adequate methods or instruments for collecting reliable evidence are available until valid and reliable techniques are evolved such objectives must be appraised by as careful subjective means as possible

4....a variety of means and techniques must be used for gathering evidence...New techniques must be developed and old techniques must be revised and modifies to meet new needs. (18)

<u>Use and Misuse of Evaluation, Measurement, and Data Obtained</u>

Many uses of evaluation have already been made evident through the discussion of the characteristics of and steps in evaluation. Shane and

McSwain and Wrightstone et al. are most thorough in their treatment of the utilization of evaluation.

Shane and McSwain distinguish 12 specific functions of evaluation in improving elementary education.

1. Focus teacher's attention on developmental changes in children as he studies their behavior and its meanings.

2. Familiarize teacher with tools and techniques used in studying and analyzing the nature and causes of behavior changes.

3. Sensitize teacher both to the subtleties and importance of interaction.

4. Increase skill in interaction as the evaluator works with others.

5. Broadens the teacher's awareness of a child in a total environmental situation.

6. By its involvement with value analysis, evaluation suggests and clarifies goals for the school and ways in which they can be captured successfully.

7. Demands constant curriculum improvement, which is in step with the changing times.

8. Provides the basis for gauging growth of a child without primary emphasis on academic achievement alone.

9. Promotes individual classroom instruction.

10. Improves public relations - involves parents and other members of the community.

11. Provides a high type of on-the-job education.

12. Strengthens democracy since evaluation depends on the use of democratic procedures for its successful fulfillment. (19)

Wrightstone et al. are more specific in their outline of the uses of evaluation data:

I. Administrative uses - provide records of pupil adjustment, interests, aptitudes, and achievement
   A. Data may be entered on pupils cumulative record and become a basis for evaluation of the individual's growth and progress or that of the class group.
   B. May provide reports to parents to supplement opinions about the

pupil by documentary evidence (tests, questionnaires, interviews, anecdotal records).

C. Make available more systematic and objective records when a pupil is transferred.

D. Provide periodic reports of school progress to patrons of the community.

E. Data from sociometric, personality, aptitude, and achievement tests maybe consulted in the classification of pupils for instructional purposes.

II. Instructional uses -

A. Supervisor may determine the status of a class or pupil in some of the major objectives of the curriculum in order to evaluate teaching methods and instructional materials, and indicate changes in instructional procedures and pupil-teacher relations.

B. Teacher and supervisor may discuss quantitative and qualitative data to agree on various instructional and learning problems.

C. Teacher may use results to determine the status of a pupil in various objectives of the curriculum.

D. Teacher may identify gifted, "normal," and slow-learning pupils.

E. Teacher may group pupils for instruction purposes within the classroom.

F. Teacher may analyze or diagnose an individual pupil's difficulty and rate of growth.

G. Teacher may determine the status of an individual or class at the beginning and end of the term.

H. Anecdotal records, observations, rating scales, personal reports, interviews, and sociometric methods aid the teacher in guiding more wisely the growth development of the pupils.

I. Data on personal and social adaptability may help:

1. identify pupils who are well adjusted and those poorly adjusted.
2. diagnose probable causes of contributing factors of the maladjustment.
3. set up individual and group conditions and situations to aid, when possible, growth toward better adjustment.

J. Sociometric data may identify leaders and isolates so that social relations may be established within a classroom that contribute to the maximum social development of each.

K. Interest inventories indicate pupil interests and permit the teacher to counsel pupils and adapt the curriculum to their needs.

L. Case studies are usually necessary for only seriously maladjusted pupils.

M. Cumulative record may be studied to evaluate and guide the growth and development of every pupil.

III. Educational and vocational guidance of most value in high school and beyond.

IV. Research uses

A. Data may be gathered by various evaluation techniques in order to study the effectiveness of different methods of teaching a subject, or meeting personal-social needs of the pupils, or in order to judge the effectiveness of an experimental program, such as the core curriculum.

B. Evaluation research may be conducted by a research bureau of the city school system, of a college or university, or an individual interested in advanced graduate work.

C. Research may be undertaken in order to obtain information pertinent in the following:

1. Diagnostic studies of learning difficulties
2. Age or grade place
3. Curves of learning for various curricular materials
4. Correlation among various measures of aptitude, ability, and personality characteristics, and case studies of particular pupils. (20)

Wrightstone and his collaborators go on to summarize misuses of the results of evaluation as being:

1. If the design of the appraisal program is narrow and limited, the testing program may tend to determine emphasis upon specific objects of the curriculum to the detriment of others and to the detriment of desirable trends in pupil growth.
2. Test results of a partial nature may be used to estimate or rate a teacher's teaching ability.

3. Reliable and valid instruments of measure are restricted to an appraisal of limited aspects of pupil behavior or growth, thus making it impossible to measure the whole result of an educational process by one test or a battery of tests. (21)

Douglas and Spitzer elaborate more fully on the effects of measurement or they point out its possible effects upon instructional procedures, learning procedures, and research. If tests, required for use by a supervisor or the administration to measure pupil achievement, measure learning outcomes which the teacher has been trying to achieve with the pupils, then he is encouraged to continue or intensify efforts in that direction. Effects can be detrimental if the teacher stresses factual knowledge in his instruction and the tests measure only this learning outcome or when a teacher is discouraged from emphasizing proper learning outcomes because tests measure only factual knowledge. Concerning the learning procedure, children learn the wisdom of learning what they will be tested on. The kind of evaluation employed affects the children's learning procedure and determines their actual learning objectives.

> Tests and activities included in other evaluation techniques give children their most tangible clues to the goals they are to achieve. Evaluation must be designed so children may identify the right goals for their learning. If understandings are our goals our practices of evaluation must include means of measuring such outcomes.

Inadequate representation of understanding in standard and other tests has affected educational research adversely. Douglass and Spitzer point out that the worth of experience program is usually appraised in terms of test data, and frequently a standardized test is the only instrument used to measure the outcomes. Even when researchers make their own test, the practice of measuring only unrelated facts and formal skills is prevalent. Data obtained in such a way may not show significant differences between the experimental and conventional procedures even when differences which are not educationally significant. Consequently, in many experiments, the evaluation has been incomplete or inappropriate. The writers are quick to add that this is not inevitable; meaningful learning outcomes can be measured. (22)

One of the important functions of evaluation, as recognized by all of the writers but especially emphasized by Rothney, is to aid in reporting.

> There is a demand for information that will tell parents and others with definiteness where their children are showing strengths and weaknesses as judged by normal expectations of children of their ages and opportunities...also a demand for information that describes a pupil's progress in a way analytical enough to give helpful guidance and to indicate pupil's likelihood of success in continued work...need for an invention of a way to direct the minds of pupils, parents, and classroom teachers away from marks toward the fundamental objectives of education. Such reports need to show appreciation for the poorest pupil's good qualities, while the best pupil's weaknesses are pointed out. When reports can do these things and also add recommendations of ways in which pupils can be helped to overcome weaknesses and use strengths more effectively, they can become potent tools in the improvement of schools. (23)

Related to Rothney's statements concerning reporting is the list of shortcomings of grades presented by Baron and Bernard. Marks tend to:

1. Become an end aim of education
2. Emphasize subject matter (pupil development should be the aim)
3. Discourage good teaching (grades become incentive)
4. Cause teacher to overlook differences, except pupils to progress at same speed, in same way
5. Create a situation that is "unlike life"
6. Penalize those pupils most in need to help
7. Have little meaning in themselves (subject to different values, judgments, interpretations, etc.)

Baron and Bernard contrast genuine evaluation and grading as 1) communication <u>with</u> versus communication <u>to</u> the person concerned, 2) competition with oneself versus competition with others, 3) emphasis on pupil development versus subject-matter emphasis, 4) guidance versus judgment, and 5) emphasis on the products of learning versus the symbols of learning. (24)

Characteristics of a Good Evaluation

Any person involved in selecting or constructing evaluation instruments or techniques needs to determine whether or not it possesses certain essential characteristics. All of the sources on evaluation which I consulted were agreed upon these characteristics with differences only in degree of emphasis or in slightly different classifications. For example, Wrightstone et al. give the qualities for judging evaluation techniques as validity, reliability, objectivity, norm, and practicability; (25) Torgerson and Adams classify them as validity, reliability, adequacy of norms, and usability; (26) and Ahman and Glock list the characteristics as norms, validity, and reliability. (27) In the latter two reliability is considered to include objectivity. Wrightstone and his associates describe the qualities for judging evaluation techniques as follows:

1. Validity - characteristic which indicates the relation of the measurement or diagnosis with meaningful criteria of learning or behavior. Some techniques are selected to show the effectiveness to predict future performance, others to indicate immediate status, establish the representative nature and scope of content or behavior, or provide data for the support or rejection of some psychological theory.
2. Reliability - indicates consistency, equivalence, or stability of the measurement that is obtained.
3. Objectivity - indicates identity or similarity of scores or diagnosis obtained from same data by equally competent scorers.
4. Norm - provides average or typical value for a measure or diagnosis obtained by administering a measuring instrument to a specific population so subsequent scores or measure of an individual or group may be compared with the typical values of a normative population.
5. Practicability - indicates feasibility for general use of the test or evaluation technique on such basis as cost, time required for administration, ease of administration, ease of scoring, ease of interpretation of the results. (28)

Torgerson and Adams go on to discuss reliability, making several statements of importance to the concern in this paper.

They point out factors affecting the reliability of any test as being the length of the test and the objectivity of the test with respect to its administration and scoring. The longer a test is the more extensive is the sampling. The more homogeneous or less variable the characteristic one is testing, the smaller the sampling can be. In selecting or constructing a diagnostic test one need not be so conscious with the length as with the number of items of each type included. As far as objectivity is concerned, it should be considered a means to an end. Emphasis on objectivity of scoring, say the writers, has led to shallowness or superficiality in measuring. The goals for which achievement is measured most easily and most objectively have been given greater emphasis in the evaluation program than other more significant objectives. They advocates as the best procedure making an evaluative instrument which achieves as high degree of objectivity as possible without sacrificing the more important criterion of validity.

Torgerson and Adams illustrate with an account of observations to remind us that the factors described as affecting the reliability of a test are equally significant for informal evaluation techniques. The sampling of behavior observed through particular procedure must be of sufficient size and representativeness. The more complex and heterogeneous the behavior being studied, the more variable it is from situation to situation and the broader the sampling should be. Objectivity in applying informal techniques is also important, but should not be gained at the cost of validity. (29)

Major Evaluation Techniques

By the reading of Ahman, Baron, Shane, Torgerson, and Wrightstone and their associates it is possible to obtain an understanding of the types, uses, and qualities of major evaluation techniques. The authors use different formats but include generally the same techniques described at one place or another in their books.

Ahman, Glock, and Wardeberg list common evaluation procedures as such:

A. Testing procedures
  1. Individual and group tests
  2. Informal and standardized tests
  3. Oral, essay and objective tests
  4. Speed, power and mastery tests
  5. Verbal, nonverbal, and performance tests

B. Non-testing procedures
  1. Interviews - a.) Teacher-pupil and b.) Teacher-parent
  2. Questionnaires
  3. Anecdotal records
  4. Sociometric devices
  5. Ranking and rating procedures

Baron and Bernard describe different techniques which are most beneficial in specific evaluations in a particular area dealt with in each individual chapter in his book. One chapter is concerned with evaluating pupil achievement and would include the most important techniques for obtaining this information; another chapter concerns studying interests and attitudes and might include some of the same techniques but would also include others of little use in evaluating achievement. (31) Torgerson and Adams (32) and Shane and McSwain (33) use this same technique while Wrightstone, Justman, and Robbins devote a whole chapter to each of their classifications of evaluative techniques, defining the techniques, describing their data and results for various uses. The techniques are categorized as:

1. Objective tests
   Classified by a) psychological characteristics and b) technical features
2. Anecdotal records and observational techniques
3. Oral and essay exams
4. Questionnaires, inventories, and interviews
5. Checklists and rating scales
6. Personal reports and projective techniques
7. Sociometric methods

8. Case study

9. Cumulative record

Wrightstone et al. deal with various evaluations, such as evaluating achievement in language arts and mathematics, evaluating interests, evaluating thinking and problem-solving, in separate chapters. In each chapter he suggests various evaluative techniques applicable to the particular evaluation desired. (34)

We notice in the classification of techniques given by Wrightstone et al. emphasis on the informal and subjective, or non-testing procedures. Great emphasis is put on these informal evaluation techniques in all sources on evaluation consulted. All the writers stress the failure of standardized and teacher-made objective tests in getting a true picture of the whole child.

Rothney points out the main purpose of testing is not to rank or grade pupils but to assist classroom teachers in obtaining evidence of growth. He is emphatic in stating that specialists in measurement have largely failed in constructing standardized tests to measure the totality of behavior: "The traditional mathematics and logic which they use are handmaiden's in the science of a piecemeal mechanistic view of the world... They have not yet produced the total understanding of the human personality which the classroom teacher must have." (35)

Wrightstone and associates state that although test technicians have recognized both the ability to apply information and skills in problem-solving and related factors, such as interests, attitudes, and personal-social adjustment, as learning outcomes, these technicians showed little concern about the measurement of the less tangible aspects of educational growth. (36)

The various writers in <u>The Mental Measurements Yearbook</u> seem to reiterate this conclusion in their evaluation of the various published measurement devices, especially those to measure achievement. (37)

Douglass and Spitzer, in 1945, showed condern over teachers in increasing numbers adopting "new-type" objective test forms, convinced that the essay test was unreliable and that its weaknesses could be avoided by using objective tests. Since framing "thought questions" in the pattern of these tests was difficult, too frequently objective testing became exclusively fact testing. The writers go on to criticize other procedures to evaluate learning pointing out that many times oral questioning can be satisfied with memorized statements and mechanical skills and that often the wrong aspect is appraised in a wide range of "work products," such as a composition or bird house, by which we undertake to judge learning. Douglass and Spitzer emphasize that tests made by teachers, other forms of evaluation, and standard tests are alike in giving little attention to understanding and call for the condition to be corrected, the effects upon teaching, learning, and research being great. (38)

Torgerson and Adams look at the differences between objective and subjective methods of evaluation and state that by subjective techniques the results vary somewhat according to the background, point of view, and the standards of judgment applied by the evaluator. The effectiveness of subjective methods depends upon the training, experience, and skill of the teacher and upon her freedom from personal prejudice or personality traits which would tend to bias or distort her results, but these methods have the advantage of permitting the teacher to study the "whole child" in a natural setting. These two men are especially concerned with our recognizing that the child reacts as a total being and that our understanding of this is essential if we are to educate or modify any part of him. Noting variations in a child's spontaneous responses to different environmental situations are especially helpful to the teacher. They emphasize the importance of the teacher developing a proficiency in using subjective methods, showing that growth in the ability to study children takes place over a period of years and must include 1.) opportunities to study individual children, 2.) an understanding of child psychology, 3.) an understanding of child growth and development, and 4.) increased objectivity. (39)

Baron and Bernard speak of anecdotal records, rating scales, and observation as providing valuable data for evaluation and goes further to suggest the use of creative writing, drawing, and painting in evaluating pupil behavior. (40) They discuss as promising practices in evaluation 1.) letters to parents, in place of the conventional report card, 2.) home visits, 3.) teacher-parent conferences in the school, 4.) self-appraisal by the pupil, 5.) teacher-pupil conferences, 6.) teacher-pupil-parent conferences, and 7.) culnulative records, including personal data, chronology and address, conference notes, record of attendance, achievement-test data, significant behavior or personality observations, and anecdotal reports. (41) Later in this paper their consideration of self-appraisal as an important educational objective and of assistance in the evaluation process will be mentioned.

Teaching For and Evaluating Understanding

Douglass and Spitzer in The Importance of Teaching for Understanding speak of understandings as practical necessities, without which we can react only to familiar situations, and then only in a routine manner. They point out advantages of learning with understanding, its economomy, permanence, cumulative effect, and functional value. The modern conception of education as no longer being concerned with closed systems and compartments of learning, but with ideas, attitudes, and skills, which influence actual living now and in the future, calls for a measurement of understanding which the writers feel is neglected. They examine samples of tests which are found to measure little else but factual knowledge. (42)

Findley reminds us that it is unreasonable to expect any teacher to measure all understandings, or even to measure all aspects of a single understanding. He lists as principles applicable to the evaluation of understanding, some of which may give us new insight:

1. In every subject-matter area there are available at present many well-known procedures for the evaluation of understanding.

2. To provide evidence of understanding, an evaluation situation must contain an element of novelty, but not too much novelty.

3. Understanding is of many kinds and many degrees and evidence is to be sought on appropriate levels.

   a. variation in subtlety and complexity of situations

   b. variation in intellectual and social distance

4. Procedures employed to measure understanding should provide evidence of appreciation of primary reality.

5. Since intelligent behavior in many situations involves the ability to recognize the relevancy and sufficiency of data, evidence of this ability should be sought.

6. Evidence of understanding is to be found in the originality of the performance on the part of the pupils.

7. Evaluative procedures should be selected with due regard for the likelihood of their evoking evidence of the kind of understanding that is required.

8. In obtaining evidence of understanding, care should be exercised to insure that a pupil's response reflects his actual level of understanding.

9. The program of evaluation should be planned so as to foster the development of habits of self-appraisal on the part of the pupils. (43)

Several of the points made by Brownell and Sims in stating the general concept and essential characteristics of understanding might be well to keep in mind when considering evaluation of such:

1. You may say a pupil understands when he is able to act, feel, or think intelligently with respect to a situation.

2. Rather than being all-or-none affairs, understandings vary in degree of definiteness and completeness.

3. The completeness of the understanding to be sought varies from situation to situation and varies in any learning situation with a number of factors.

4. Pupils must develop a worthwhile understanding of the world in which we live as well as with symbols associated with the world.

5. Most understandings should be verbalized, but verbalizations maybe relatively devoid of meaning.

6. Understanding is developed as a pupil engages in a variety of experiences rather than through doing the same thing over and over.

7. Successgul understanding comes in large part as a result of the methods employed by the teacher.

8. The kind and degree of a pupil's understanding is inferred from observing what he says and does with respect to his needs. (44)

Evaluating Attitudes, Values, and Interests

Wrightstone and his collaborators devote a chapter to the consideration of evaluating attitudes and values. He looks at the various techniques which may be used in this evaluation. The few standardized techniques have limited application in the classroom and their curriculum relevance is relatively low. Teacher-made techniques include observation and paper-pencil tests, anecdotal records, checklists, and logs; the presentation of a variety of stimuli and of sociodrama to elicit specific behavior for observational purposes. Paper-and-pencil tests, the most frequently used in attitude testing, include agree-disagree inventories and check-lists, rating scales, social-distance, paired comparisons, and free-response approaches. Wrightstone emphasizes that a favorable attitude toward a person, object or activity, connected with school is more likely to motivate a person to do well, and since negative attitudes toward these factors hamper maximum learning, the teacher should be aware of attitudes. Attitudes and values also indicate how one can expect people to behave in future situations. The importance of attitudes and values should be recognized in the whole process of problem-solving. (45)

Wrightstone attempts to define attitude and values and further states that they should not be seen as separate and distinct from such elements as interests, needs, aptitudes, and emotions, but as interrelated factors in the nature and development of the total person. In a learning situation, new information and knowledge may lead to a change in attitude and values; the development of proficiency in any given skill may result in favorable attitude toward participation in a particular classroom activity. (45)

Baron and Bernard speak of studying interests and attitudes in the classroom as providing information useful in 1.) understanding pupils, 2.) discovering motivation possibilities, 3.) relating teaching to pupil's interest and experiences, 4.) studying and evaluating pupil's interest changes, 5.) helping pupils to (a) become aware of their interests, (b) evaluate their interests, and (c) increase their understanding of themselves, and 6.) stimulating thought and discussion among pupils concerning the implications of their interests. (46)

Baron and Bernard list as methods of studying interests observation, interviews, written reports and tests of pupils, questionnaires, and specially designed inventories. Listed methods for studying attitudes include observation, interviews, ratings, various types of attitude and opinion scales. Points concerning observational records are suggested. (47)

Evaluating Thinking and Problem-Solving

Another chapter in Evaluation in Modern Education concerns the evaluation of thinking and problem solving. John Dewey is referred to in stating that the complete act of thinking involves the task of finding meanings and trying them out in order to determine whether a particular course of action satisfactorily solves a problem in accord with social values. Thinking is the process of solving problems. Wrightstone et al. view the problem-solving situation as a complexity of anumber of elements, consisting of;

1. A problem (What should I do?)
2. Courses of action (Which course of action should I take?)
3. Limitations and assumptions (What conditions do I take for granted?)
4. A set of values (What actions will be consistent with my values?)
5. Reasons and consequences (What will happen if I do each of these things?)
6. Solution and verification (What best solves my problem?)

There are few standardized tests for evaluating pupil thinking and problem solving ability. Observation and interview are most useful in all grades, and paper-and-pencil techniques may also be used in upper grades. (48)

Educational Diagnosis

Of assistance to us in contemplating evaluation may be Torgerson and Adam's scrutiny of measures of general status and diagnostic measure. A measure of general status is one in which scores are used as a basis for marking, grading, or ranking pupils, for determining how much of a skill, ability, or personality trait a pupil possesses. In a diagnostic measure the total score is relatively unimportant, and the pattern of response is of major concern. Distinctions are not always clear-cut; some tests, such as the California Achievement Tests, are tests of general status, but many features assist in diagnosis.

The writers review the questions to which diagnosis can provide answers. Results from diagnostic tests assist in answering "Who are the pupils having trouble?" and "Where are the errors located?" Additional diagnosis is required to provide more specific answers to the second question and to reach the third level of diagnosis, "Why do these errors occur?" (49)

Later in their book a chapter is given to the treatment of educational diagnosis, including its importance, the level of diagnosis, steps in the process, and basic principles of corrective instruction. (50)

Our Concern in Evaluation

Before concluding this section concerning evaluation in general education I would like to mention that Torgerson's and Adam's consideration of the place of evaluation in the teacher's program may prove very useful by pointing out certain limitations, taking into consideration her total responsibility involving the group instruction program and the study of the individual child and by emphasizing the impossiblity of effective teaching without studying and understanding each child. (51)

As we prepare to focus on the evaluation in mathematics it will serve us to keep several summarizing points given by Torgerson and Adams concerning evaluation and the modern school.

1. The widespread individual differences in ability and in achievement in a typical class pose a crucial problem which the teacher must understand and resolve.
2. As significant as the variability in aptitude and attainment within a class is the variability of strengths and weaknesses within each individual child.
3. Studying and understanding the children are essential to good teaching.
4. Standard tests are important tools to aid teachers in understanding children.
5. The use of standard tests needs to be supplemented by other techniques of child study.
6. Many techniques of child study must be understood and their uses, limitations, and advantages recognized.
7. Child study data must be interpretted objectively.
8. Measurement, evaluation and individual instruction are interrelated aspects of effective teaching.
9. The school must accept the responsibility for teaching every child how to read, compute, write, spell, and acquire understanding within the limits of his capacity.
10. An effective program of individual instruction reduces the need for remedial education. (52)

RELATED READINGS

IN THE AREA OF

EVALUATION IN MATHEMATICS

READINGS IN THE AREA OF EVALUATION IN MATHEMATICS

Change in the Concept of Mathematics

The changes which have taken place in the concept of education, as enumerated in the preceding section, should perhaps have served to revolutionize instruction and evaluation of mathematics more than some other subject areas. Let us consider what several writers have said concerning a change in mathematics instruction and evaluation.

Bathurst, in How Children Use Arithmetic, states three major objectives in helping children use arithmetic today:

1. To help children understand number concepts in daily living.
2. To teach children to compute or figure with the skill required in everyday activities.
3. To help children improve their ability to recognize in their life problem situations that require computation or other use number.

Bathurst stresses that a large part of the emphasis on arithmetic in modern schools is concerned with helping children understand the system of arithmetic or number which we use and its place in our culture. Instruction begins (1) when children have a problem such as finding how many books are needed for a small group and (2) when they have had experience counting objects and reasoning about them and adding and subtracting through moving objects about. (53)

Buswell speaks of the conspicuous gain in the last 15 years as being the general acceptance of the idea that arithmetic must be taught meaningfully. He goes on to discuss the difficulty of discovering how pupils think when the outcome is understanding. The main emphasis of this writing is that we need to obtain more knowledge of the way in which pupils think at various levels of learning. He lists six methods of studying pupils thinking, proposing that more attention be given to descriptive studies. (54)

Weaver writes, in 1955, that the change in arithmetic instruction in the past two decades has brought increasing concern for the process of learning as contrasted with the product of learning. How children learn is at times more important than what they learn. He re-emphasizes points made by Buswell concerning the teacher's awareness of the levels of thinking employed by her children and the importance of the use of the interview technique in which children "think aloud" as they respond to specific quantitative situations and their related questions. (55)

Miller analyzes the effectiveness of the meaning method and concludes through his experimentation that the meaning method is definitely more effective for the area of computation of fractions, in establishing retention in the process of computation as well as in the understanding of the principles of arithmetic, and for the comprehension of complex analysis in arithmetic indicating a potential superiority for difficult concepts. (56)

Glennon in Testing Meanings considers the lag in the development of adequate methods and devices for measuring growth in understandings and meanings in arithmetic and attributes certain causes to lags in other aspects of the arithmetic teaching-learning situation. (57)

Ahman et al. speak of technology and the jobs of the physicist, banker, astronomer, and economist, the inconceivable speed and accuracy of the calculations of computors. The goals of instruction reflect these trends, say the writers; no longer is teaching "the tables," and the "goes-in-to's" an adequate goal for teachers. Understanding is the important outcome. (58)

Another recognition of change is noted in the writing of Shane and McSwain. They advise that the process of evaluating the results of teaching and learning in the meanings approach requires professional competence on the part of the teachers, more class time devoted to questioning, reasoning, and proof, and greater emphasis on teacher diagnosis and children's diagnosis of how problems are solved. (59)

The current trend for math instruction to stress the development of understandings of math principles, generalizations, and relationships than had previously been done is re-emphasized in Gray's and Moody's writings in 1966 and in 1968 by Ashlock. Gray points out that arithmetic facts and computational skills are now seen more as tools for the developing deeper and richer insights into the nature of math than as ends in themselves. More elementary pupils are encouraged to think and act like mathematicians, discovering important mathematical ideas for themselves. There is refocused attention on the problems of evaluation. Gray reasons that the traditional evaluations of pupil progress by instruments which primarily were measures of speed and accuracy were defensible while arithmetic was conceived as a body of facts and computational skills but are inadequate as measures of the kinds of understanding now stressed. Feeling an urgent need for an instrument to evaluate both pupil progress toward acquiring these mathematical understandings and the effectiveness of the new programs of instruction, Gray describes his experimentation with an interview inventory in "An Approach to Evaluating Arithmetic Understandings." (60)

Moody states that it is desired that a student develop a maturity in thinking mathematically and in facing new mathematical situations. It is now seen as important that a student understand the structure of mathematics and develop an idea of how a mathematician thinks. There is an attempt to wean students away from accepting statements without questioning and from overgeneralizing and assuming. He suggests a test which attempts to measure a person's ability to handle mathematical concepts. (61)

Ashlock recalls that the findings of psychological and educational research support the emphasis placed on mathematical understanding in the past 30 years. He feels that the existing literature indicates that arithmetic achievement tests have been made to create tests of the understanding of the properties of the number system. The purpose of his study is to develop such a test. (62)

Sueltz enumerates recent developments and pressures which have caused evaluation of the instructional program in mathematics to become more important.

1. New mathematics curricula are being advocated and tested by experimentation and research.
2. New mathematics content is available and is being proposed for inclusion at several levels of instruction.
3. New devices and materials of instruction are being developed

and promoted by financial resources never before available to our schools.

4. New principles of learning are being emphasized in the presentation of mathematical concepts.
5. Society is demanding greater mathematical competence of all citizens than ever before.
6. National survival may depend upon the development of new mathematical concepts. (63)

The Program of Mathematics Evaluation

The definitions, characteristics, and principles of evaluation provided in the preceding chapter are helpful to us in determining of what mathematics evaluation should consist. Let us now focus our attention on writing pertainly directly to the subject of evaluation in mathematics.

Evaluation is distinguished by Grossnickle and Brueckner as a "continuous process of inquiry concerned with the appraisal, study, and improvement of all aspects of the instructional program in mathematics, including pupil achievement." The evaluation process is broken down into six basic steps which are briefly explained by the writers.

1. All major goals and values of the mathematics program must be determined and accepted. These reflect the ideals and wishes of the community.

2. The objectives, both immediate and ultimate, should be based on a systematic analysis of individual and group needs,...clarified and formulated in terms of desirable behavior on the part of individuals and groups concerned.

3. Steps must then be taken by appropriate procedures to collect evidence of achievements and growth with respect to the established goals and values as revealed by changes in the behavior of the learners in social situations in the work of the school and community life.

4. There should be an examination of the school environment and of instructional practices...The contacts of children both in and out of school should be analyzed.

5. Synthesis and interpretation of all of these findings concerning pupil growth and educational practices is the final step in evaluation,

leading to redefinitions of goals and values, as may appear desirable, and to the planning of improved ways and means to attain the accepted objectives.

6. The schools should act throughout to secure the interest and cooperation of parents and all community agencies concerned with growth and development of children in evaluating the total mathematics program and in planning its improvement. (64)

Grossnickle and Brueckner give five basic steps in developing methods of evaluating the important outcomes of the mathematics program that are helpful in selecting and constructing appraisal instruments.

1. Formulate the objectives clearly
2. Clarify the objectives
3. Collect test situations and items
4. Record the behavior
5. Evaluate the behavior and interpret the record (65)

Hartung begins his chapter on "Basic Principles of Evaluation" with a description of the nature of the evaluation process. After definition of various terms he finally defines evaluation as the process of finding the extent to which the actual experiences (behavior in specified situations) conform to the objectives (desired experience). He notes the distinction between evaluation and measurement.

> In education the term 'measurement' refers to the process of gathering certain types of data for the purpose of evaluation. It also includes part of the process of analysis and interpretation, but not all of it. The calculation of the mean score for a class that has taken a test is part of the measurement process. The interpretation of this score will usually carry one into the broader domain of evaluation. Interpretation may require value judgments that are based on the value position or educational philosophy of the teacher or school. It is this aspect of the process which has led to the adoption, during the 20-year period following 1934, of the term 'evaluation' for the complete process.

Appraisal, according to Hartung, refers to a comparison of the measured achievement of one group of students with the achievement of other, presumably comparable groups. (66)

Basically, states Hartung, the purpose of education is to change students from a given state of experience to a desired state by means of a variety of appropriate learning experiences, some of which may be used as a basis for evaluation of achievement. To evaluate in this way, we measure at intervlas and try to determine whether any change has occurred in the students.

Hartung gives the steps in the evaluation process as follows:

1. Adequate formulation of the objectives for evaluation purposes
2. Selection of sample situations in which the desired student behavior may legitimately be expected
3. Obtaining a record of what happens in the sample situations
4. Analysis and interpretation of the record
5. Determining desirable changes in objectives, learning activities and instructional methods, techniques of evaluation, materials, and facilities, and other aspects of the curriculum (67)

Principles for formulating objectives, measuring attainment of an objective, and the criteria for stating objectives are treated fully by Hartung before he continues to speak specifically about the program of evaluation. According to Hartung, the evaluation program refers to all the activity of the school staff in collecting and interpretting data for evaluation purposes. The program should satisfy certain criteria, including:

1. Comprehensiveness. The program should try to evaluate with respect to all the important objectives that are identified by the teachers.
2. Balance. The emphasis on all the important objectives should be proportionate.
3. System. Evaluation should take place at regular intervals so that growth may be measured.
4. Changes effected by evaluation (68)

Brownell re-emphasizes that there should not be a separation between measurement and instruction. Attempts to establish a separtion has unfortunate effects upon classroom evaluation and teaching alike, such as 1.) removing the measure further from the immediate learning situation, 2.) limiting the measurement to outcomes most easily assessed, 3.) limiting unduly the techniques serviceable for evaluation, and 4.) creating confusion with respect to purposes of evaluation. (69)

Of worth for us to keep in mind are several recommendations offered by Brownell 1.) Evaluation and teaching should start with all arithmetic outcomes which are deemed worthy of attainment; evidences would then be obtained, as far as possible, on all types of growth at all stages in this growth. 2.) Any procedure which might shed light on learning should be accepted and utilized. 3.) Evaluation procedures should be adapted to the ends for which they are best fitted and 4.) evaluation should be continuous with teaching and learning, with evidence being collected daily. 5.) Evaluations should be immediate and intimate, reflecting the unique conditions, emphases, and factors affecting learning in particular classrooms. (70)

Brownell views comprehensive and functional evaluation in arithmetic as dependent upon the effective relating of a number of factors.

1. An adequate statement of arithmetic outcomes
2. Recognition of the peculiar demands of different purposes in evaluating learning as the demands affect evaluation procedures and instruments
3. A practicable program of evaluation
    a. knowledge of effective procedures and their limitations
    b. a realistic understanding of what can and cannot be done by the classroom teacher
    c. the actual planning and designing of the evaluation procedures (71)

Marks, Purdy, and Kinney include an extensive discussion of planning and carrying out the evaluation program. They remind us that the emphasis in any given program is determined by the purposes. For instance, for diagnostic purposes, a highly detailed analysis is occasionally required in some outcomes. Of note is their comment that any outcome that is neglected in the evaluation is inevitably overlooked in the learning activities.

Marks et al. restate other writers description of measuring the effectiveness of learning procedures by the extent to which learning has occurred and the need for teachers to continuously appraise their own success. It is necessary to determine not only how plans are working in general but the effect on each individual pupil. The writers stress that for this reason all classroom activities must be conducted on an experimental basis and continuously adjusted to the process revealed by appraisal.

The planning of the program of evaluation requires further definition of each outcome in terms of pupil behavior. Procedures are then required for providing situations in which the behavior will be revealed, and for collecting a record of the behavior. Regardless of the purpose of evaluation and the procedure say Marks et al., the key question is, "If the procedure were effective what results would be observed?" and more specifically, "What kinds of pupil behavior would indicate learning, or failure to learn? Where, and in what situation will such behavior occur? How can a record of such behavior be obtained?" (72)

The Role Evaluation May Play

Grossnickle and Brueckner explain that on the basis of the information about children's growth and development secured by evaluation procedures, judgments can be made as to the effectiveness with which the mathematics program meets the needs of the children and of the community and also about the quality of the program. Subsequently, plans for dealing with the problems that emerge can be considered. (73)

Brownell catergorizes the purposes of evaluation as 1.) diagnosing class and individual difficulties, 2.) inventorying knowledge and abilities, 3.) determining the extent of learning over a limited period, 4.) measuring learning over a relatively long period, and 5.) obtaining a rough measure for comparative purposes. (74)

The purposes of the evaluation program, according to Marks et al., besides providing a continuous check on instructional procedures, are as follows:

1. To reveal stages at which pupils have arrived in the learning process.
2. To identify group and individual need for review and reteaching.
3. To record pupil progress, and report it to pupils and parents. (75)

Sueltz's chapter in Evaluation in Mathematics discusses the role of evaluation. Several excerpts may prove enlightening in showing ways in which evaluation may be used.

> Evaluation, an essential part of the mathematics program at every level, should be the handmaiden of instruction and learning. It is not a separate entity in a good school program. It may serve to improve the instructional program in the school, to enhance the effectiveness of the teacher, to aid the student in learning mathematics, and to furnish valid data for research.
>
> Evaluation can identify and define steps and levels in the sequence that are appropriate for a given grade or age level. Careful evaluation should show not only how far a pupil has progressed in

the major steps of a sequence, but also how well he has understood and mastered a particular step. Good evaluation will show the facts and skills mastered (and those not mastered) by the student, his attitude toward the subject, and the depth of understanding and insight accompanying his work.

Evaluation is useful in determining the relative use or difficulty of learning, applying, or remembering a topic, and it serves as a guide in scaling and arranging curriculum materials.

Evaluation provides valuable data that is used in reporting to parents, in advising on school programs, and in studying many aspects of curriculum and school programs.

Measurement can determine a great deal about the student's mental ability, his likes and dislikes, his social development, his emotional maturity, his determination, his health, and his background of experience and learning at a particular school level. This information has an immediate bearing upon his potential success in a given area of mathematics.

When a student takes a test he must think and perform operations, usually with a feeling of concern and intense concentration. His test response, therefore, are apt to be remembered longer than a casual learning experience. Quizzing informally or interviewing for evaluation purposes may also be used to foster reorganized thinking and better understanding. (76)

Criteria for Judging Evaluation Techniques

The various writers emphasize the importance of carefully scrutinizing the qualities of any method or instrument of evaluation. The variety of classifications and emphases may be great but most evaluators use the same general criteria.

Hartung discusses means of securing a record of student behavior, naming validity, relevence, reliability, and objectivity as necessary qualities of a measurement techniques. (77) The criteria Grossnickle and Brueckner designate to be considered in selecting appraisal methods are similar but include a helpful explanation of each.

1. The Values of the Characteristic or Aspect of Growth Tested

   Does the test claim to measure an aspect or characteristic of pupil growth in which you are interested? Are the educational outcomes listed of undoubted value and significance?

2. Validity

Does the test actually measure what it purports to measure?

3. Reliability

Does the test measure accurately? Will children tend to get the same scores if measured again?

4. Ease of Administration and Scoring

Are the directions for administering the test clear and easy to follow? Is the test fairly easy to score? Are the tabulation forms clear?

5. Provision and Usability of Standardized Norms

Does the test yield scores that are well defined and adequately standardized? Are the scores readily understandable? (78)

What to Evaluate

As we contemplate what to evaluate in mathematics it is important for us to keep in mind that the teacher of arithmetic is the teacher of the whole child, as Glennon and Callahan state, and it is his responsibility to develop those social and mathematical understandings, attitudes, habits, appreciations, skills, and abilities, personal, social, and moral values essential for living in a democratic society. The teacher must be aware of the total growth of the learner if he expects to guide total growth. Accepting responsibility for developing the whole child, the teacher must also accept responsibility for measuring the whole child. Glennon and Callahan review the incompleteness of a testing program using teacher-made and paper-and-pencil tests only and advocate the use of a variety of techniques and devices, allowing the study of a child's process of learning as well as the outcomes of learning.

A complete program in evaluation, say the writers, will measure growth in the ability to make judgments in quantitative situations, the ability to do mental arithmetic, attitudes toward arithmetic, the appreciation of the uses of arithmetic, and other outcomes. In addition to teacher-made, paper-and-pencil tests, and interviews, it will make use of real and contrived problems situation tests, dramatizations, anecdotal records, growth charts, and others. (79)

Ahman et al. consider the usefulness in evaluating pupil achievement of informal or standardized paper-and-pencil tests in checking whether a pupil knows the proper procedure and can make correct calculations. They also emphasize the equal importance of observing the times and the ways in which a pupil uses these skills in his personal and social life. The writers describe the best situation in which to evaluate the degree to which objectives concerning mathematics have been achieved as being a "permissive but provocative environment," a classroom in which opportunity to apply skills learned in arithmetic exists in numerous and varied forms.

Ahman and his collaborators categorize the mathematics learning which should be evaluated as follows:

1. Ability to read terms and symbols
   a. Reading symbols and abbreviations
   b. Reading graphs
2. Skill in computation
   a. Basic facts
   b. Understanding the processes
   c. Mental arithmetic
   d. Working with fractions
   e. Denominate numbers
   f. Standardized tests
   g. Estimating and checking
3. Arithmetic problem-solving ability
   a. Seeing relationships
   b. Solving problems
4. Social applications of mathematics
   a. Personal use
   b. History of number
   c. Attitude and interest (80)

Shane and McSwain pose six questions to which teachers should direct their attention when seeking information about children's understanding of and developmental abilities in arithmetic. They later treat the evaluation of desired outcomes in arithmetic by listing questions which may be used by teachers in measuring the developmental progress made by children in learning basic meanings and operations in elementary math. These questions are listed under the categories of desired outcomes a.) in counting and the number system, b.) in addition, c.) subtraction of, d.) multiplication of, e.) division of integers, fractions, and decimals, f.) related to equivalents, g.) in practical measurement, and h.) in solving problems. (81)

Many lists of objectives or desired outcomes in arithmetic learning have been recommended. Sobel, and Johnson (82) and Shane and McSwain (83) are suggested for reference, and outcomes suggested for evaluation by Grossnickle and Brueckner are presented in relation to evaluation techniques in the following section. Because of its comprehensiveness, Brownell's list of desired arithmetic outcomes is reproduced:

1. Computional Skill

    Facility and accuracy in operations with whole numbers, common fractions, decimals, and percents (This group is separated from the second and third groups because it can be isolated for measurement. In the separation much is lost; computation without understanding when, as well as how, to compute is an empty skill. Computation is only important as it contributes to social ends.)

2. Mathematical Understandings

    a.) Meaningful concepts of the quantitative, the number system, whole numbers, common fractions, decimals, percents, measurement, etc.

    b.) Meaningful vocabulary of useful technical terms of arithmetic which designate quantitative ideas and the relationships between them.

    c.) Grasp of important arithmetic generalizations

    d.) Understanding of the meaning and mathematical functions of the fundamental operations.

    e.) Understanding of the meanings of measures and of measurement as a process.

    f.) Understanding of important arithmetic relationships, such as those which function in reasonably sound estimation and approximation, in accurate checking, and in ingenious and resourceful solutions

    g.) Some understanding of the rational principles which govern number relations and computational procedures.

3. Sensitiveness to Number in Social Situations and the Habit of Using Number Effectively in Such Situations

    a.) Vocabulary of selected quantitative terms of common usage b.) Knowledge of selected business practices and other economic applications of number

c.) Ability to use and interpret graphs, simple statistics, and tabular presentations of quantitative data

d.) Awareness of the usefulness of the quantitative and number in dealing with many aspects of life

e.) Tendency to sense the quantitative as part of normal experience, including vicarious experience, as in reading, in observation, and projected activity and imaginative thinking

f.) Ability to make (and the habit of making) sound judgments with respect to practice quantitative problems

g.) Disposition to extend one's sensitiveness to the quantitative as this occurs socially and to improve and extend one's ability to deal effectively with the quantitative when so encountered or discovered.

Throughout the readings much emphasis is given to evaluation of meanings and understandings in arithmetic and ofthe child's thought process; treating these subjects most thoroughly are the writings of Ashlock (85), Glennon (86), Sueltz (87), Brownell (88), and Sueltz et al. (89). Particularly helpful on the subject of evaluating children's thinking and how they learn are the studies of Moody (90), Buswell (91), Weaver (92), Grafft and Ruddell (93). The evaluation of attitudes and appreciations is well covered in the writing of Corcoran and Gibb (94), Sueltz (95), and Dutton (96). The importance of evaluating motivation is best described by Sueltz (95).

Methods of Evaluation in Mathematics

There are an increasing number of techniques which have been used in evaluation, an increasing number which are being proposed for use, and a great need to apply these and new ideas in evaluative techniques effectively. Let us first consider general procedures in mathematics evaluation and then attend to some criticisms of present techniques and recommendations for specific evaluative procedures, paying particular attention to the less conventional, more informal methods.

The most valuable methods for evaluating mathematics learning objectives are listed by Grossnickle and Brueckner as:

1. Standard tests and objective test procedures
    a. Standard tests
        (1.) Achievement tests
        (2.) Readiness tests
        (3.) Diagnostic tests, dealing with specific phases
    b. Unstandardized, short-answer objective tests
        (1.) Simple recall or free response
        (2.) Alternate response
        (3.) Multiple choice
        (4.) Completion
        (5.) Matching
2. Evaluation by less formal procedures
    a. Analysis of behavior in some problematic situation
    b. Use of behavior records
        (1.) Controlled conditions involving checklists, rating scales, time studies, recordings
        (2.) Uncontrolled conditions involving anecdotal records, diaries, reports by self and others, observation of behavior in the classroom and elsewhere, records of social agencies
    c. Inventories and questionnaires about attitude, interest, activities, and methods of study
    d. Interviews, conferences, personal reports
    e. Analysis of qualities and merits of some product (i.e. a graph or an oral report)
    f. Sociometric procedures to study social relations (97)

They discuss standard tests as providing a basis for determining the effectiveness of an arithmetic program as a whole. By careful selection it is possible to secure fairly satisfactory measures of such outcomes as (a) computational skill, (b) ability to solve verbal problems, (c) knowledge of arithmetic vocabulary, (d) knowledge of social applications of arithmetic, (e) ability to read maps, graphs, and tables, (f) functional quantitative thinking, and (g) understanding of the number system and number operations.

Grossnickle and Brueckner demonstrate how different techniques may be applied to evaluate specific learning outcomes in the following table.

Outcomes	Evaluation Techniques to Apply
The learner is:	
1. Developing meaningful concepts of numbers and the decimal number system a.) Understands the meaning and function of place value b.) Uses symbols to express numbers of all kinds c.) Understands why numbers "behave as they do"	Objective tests of understandings Observation of daily work Interview with the learner Anecdotal records about contributions Demonstration by the learner
2. Becoming skillful in the fundamental operations and the ability to apply them a.) Has control of knowledge of the basic facts b.) Understands the meaning of the four number operations and their interrelationships c.) Has skill in computations Can solve real and vicarious problems	Standard tests Informal tests, from text or teacher-prepared Observation of behavior Analysis of daily written work Interviews to test understanding Anecdotal records
3. Developing competence in utilizing the system and instruments of measurement and quantitative procedures in dealing with problems of every day living a.) Can read and use the ruler b.) Has skill in using measures to describe and define quantitative aspects of objects, events, and ideas	Problem-situation tests Objective tests Behavior records and ratings Rating of performances product Interview Reports of responses in other curricular areas
4. Developing desirable interests and attitudes toward mathematics a.) Makes voluntary contributions of significance to class discussion b.) Reads widely about mathematics and its uses c.) Is resourceful and inventive in dealing with quantitative aspects of problem situations	Interest inventory Ratings of interest in activities and toward curriculum content Observation of behavior Self-rating devices Interview Questionnaire of activities

5. Developing effective methods of studying and learning mathematics	Tests of effectiveness of learning procedures
a.) Makes an agressive attack on learning of facts and operations	Observation of behavior Rating of behavior Interviews
b.) Uses manipulative and visual aids to develop insight	Records of activities Questionnaires
c.) Practices systematically to develop mastery of the skills	Self-rating procedures
6. Developing desirable behavior patterns and good citizen traits as aresult of group activities	Observation of behavior Problem-situation tests
a.) Reveals qualities of leadership	"Guess-who" tests Rating scales
b.) Participates effectively in group work and committee assignments	Interviews "What would you do" tests Tape recordings
c.) Is able to attack real problems systematically and effectively	(98)

Dumas suggests several ways in which we can study children to find out about individual abilities and learnings:

1. Examine records accumulated from reports of past teachers, nurse, parents, and others
2. Listen to children as they take part in discussions involving arithmetic, either directly or incidental to some other topic
3. Observe pupils at work on arithmetic problems
4. Analyze papers on which children have worked problems involving number relations. Look for maturity of method, as well as for accuracy.
5. Interview each child to discuss with him his own analysis of strengths, weaknesses, and interests
6. Give tests of various kinds
   a.) Readiness tests - discover how well prepared children are for new work
   b.) Diagnostic tests that reveal particular types of difficulties individuals may be facing
   c.) Achievement tests - indicate general levels children have reached

d.) Teacher-made tests - ascertain the success children have had with what has been taught

e.) Intelligence tests - general (and tentative) idea of each child's ability to learn arithmetic

Dumas also reminds us that the simple procedure of having a child say aloud what he thinks as he works will pinpoint his difficulty in working processes such as 2 figure multiplication, long division, or subtraction of unlike fractions. (99)

Many different classifications of evaluative procedures are given such as Brownell's and Shane's. Brownell's general classes are 1.) paper-and-pencil tests, 2.) teacher observation, 3.) individual interviews and conferences with pupils, and 4.) pupil reports, projects, and the like. (100) Shane and Brownell list instruments for and procedures in evaluation as 1.) achievement tests, 2.) informal teacher-prepared diagnostic instruments, 3.) teacher-child interviews, 4.) evaluation of written work, and 5.) use of textbooks and workbooks. (101)

Considering the emphasis on the less formal, and certainly un-standardized, techniques for evaluation it is well for us to scrutinize these techniques. Although each source contains some discussion of each of these procedures and some of them will be further elaborated upon later in this paper, we will consider the writing of Marks, Purdy, and Kinney concerning observational procedures, anecdotal records, checklists, and rating scales.

Observational procedures are possible through the teachers daily contacts with her pupils. By continuously collecting information on the progress of individual pupils, observational procedures can perhaps provide the most valuable information. Two possible weaknesses of such procedures are that 1.) there is apt to be no record available when the teacher needs it and 2.) important behavior is apt to be overlooked or forgotten. Standard procedures have been designed to make observation more valid by providing safeguards against these two possibilities, the most common being anecdotal records, checklists, and rating scales.

Marks et al. suggest that a teacher have a pad handy on the desk for keeping a record of happenings. Their design for anecdotal records provides for two entries: 1.) what happened, and 2.) what it probably means (its significance). Any interpretation of behavior is tentative and subject to revision as more anecdotes are collected. With practice a teacher can become skillful in the use of anecdotal records to make observation an effective evaluative procedure. After becoming accustomed to their use, the teacher tends to see much more of what is going on.

A checklist is defined by these authors as memoranda of behavior to look for, with provision for making notes. He gives an example of a checklist to evaluate interests of a class in mathematics with questions such as "Are they willing to do work beyond the assignment?" and suggests that with little variation, a separate blank for each pupil, a checklist could be devised to apply to each individual pupil. The notation would then include the date and concise memoranda of what happened. By looking through the blanks at the end of the month, the teacher would know who had demonstrated interest (as defined) and who had not. Whether or not to devise a checklist for a particular outcome depends upon whether there is a need for specific information. If a teacher feels that he may be neglecting certain outcomes, or if he has any questions as to the effectiveness of his procedures, then a checklist serves to indicate these points.

Rating scales may be used instead of checklists in order to record appraisal along with observations of behavior. A sample item on a rating scale might look as follows:

	Never	Occasionally	Frequently
1. Does work beyond requirements	________	________	________

Or it might be in graphic form:

1. Does not turn in assignment	Turns in assignment if urged	Turns in assignment on own accord	Does work beyond assignment

A check is place above the phase most nearly descriptive of the typical performance.

Marks et al. point out that any special observational procedures are time consuming; whether any should be used depends on whether information is really needed and whether the outcome is important.

They continue to take under consideration teacher-made tests, examining types of questions, diagnosis of pupil difficulties, and standardized achievement tests. (102)

Some criticisms of some evaluation procedures and recommendations for new techniques have already been evidenced. Glennon in Testing Meanings in Arithmetic relates the causes of the lag in the development of adequate methods and devices for measuring growth in understanding and meanings in arithmetic to the lag in other aspects of the arithmetic teaching-learning situation.

The causes are stated as:

1. The changing role of arithmetic in the curriculum from the science of numbers to a "tool subject."
2. The narrow interpretation of practical education (philosophy of pragmation) well supplemented by the psychology of the teaching-learning situation (emphasizing facts and processes).
3. The mental security which teachers and supervisors find in the present process of telling, drilling, testing.
4. Thinking of arithmetic as aseries of arbitrary associations, each association being an entity in itself and having no relation to other associations.
5. The degree to which presently available tests impinge upon and striate the aims and objectives of learning in arithmetic. (Teachers tend to teach for learning they know is included in standard test.)
6. The lack of a definitive list of understandings and meanings. (Lists are ambiguous.) (103)

This particular writing of Glennon concerns his frontier research study in the area of testing for meanings in arithmetic. His method is the mass administration of a paper-and-pencil test and the data gathered provides limited evidence of the meager degree to which teachers up to that time had brought about growth in meanings and understanding. Glennon emphasized the need for more such studies to help teachers in teaching and testing for meanings in arithmetic.

Buswell's research mentioned earlier proposes that methods other than objective need to be used in seeking an understanding of how pupils think when working with numbers. He suggests as supplementary to some of the objective studies, more attention be given to descriptive studies of how pupils think. Recognizing that these methods leave something to be desired in the way of quantitative evidence, he feels they may provide leads which later may be followed by more objective methods. He emphasizes the difficulty in discovering how pupils think when the outcome is "understanding rather than confusion." The point he brings out through his writing is that we do not know the nature of the meaningful experiences which make up the thinking of pupils through various stages of learning until a full maturity of understanding is reached. He presents six methods of studying pupils thinking.

1. Teachers study the development of meanings from year to year as pupils progress through the grades.
   a.) Keep a systematic record of pupils' responses to a common group of exercises or questions
   b.) Statement of recommendations which indicate the nature of a pupil's understanding from grade to grade and shed light on the problem of learning at various levels of maturity
2. Keeping records of a pupil's responses based on "thinking aloud" while he carries on arithmetical operations.
3. Providing pupils with an ample supply of manipulative aids and asking them to illustrate correctly algorisms they express abstractly in numbers.
4. Making a diagram of various procedures in solving a problem and use this as a diagnostic tool, tracing the steps of an individual's thinking.
5. Use of unusual procedures in computation and the use of number systems with basis other than ten (Helps one become more aware of thinking procedures involved in novel situations).
6. Methods of getting objective evidence.
   (example: using a metrome or string pendulum to make possible a time analysis of a pupil's operation in arithmetic) (104)

In speaking of progressive levels of meanings Buswell explains that the same operation has different meanings to a child as his understanding of the number process develops. Teachers need to know both the array of mathematical meanings which make-up a subject and the levels at which various meanings can appropriately be learned. The teacher must understand the successive levels of a pupil's thinking in order to help him develop clear meanings at every stage of the learning process. (105)

Weaver elaborates on one of Buswell's techniques involving "thinking aloud." He illustrates with a description of a particular teacher's procedure of presenting individual children with 3" x 5" cards each containing a different multiplication combination. If the child's answer to each was not automatic then he "thought out loud" as he attempted to find the product, and the steps in his thinking were recorded. (106)

Gray examines the interview technique describing it as facing a child with a problem or example, letting him find a solution, and then challenging or questioning him to elicit his highest level of understanding of the process. Gray directs attention to weaknesses of the interview technique which have left findings of such techniques open to question.

1. Reliability
   a.) Test-retest reliability - There are only a small number of items and because of the practice effect test-retest reliability is doubtful. Reliability must rest upon careful interview procedures and the training of the interviewer.
   b.) Inter-interview reliability - When a team of interviewers are used it is desirable to be assured that all members of the team will score a given response the same way.
2. Economy of administration - Some interviews require 45 to 50 minutes.
3. Statistical analysis - Data from interviews are frequently reported in terms of gross results or in percnetages which do not lend to sophisticated statistical analysis. This is due partly to the fact that such statistical procedures are not widely used and partly to a lack of sufficient discrete categorization of interview responses. (107)

Gray attempted to develop an individual interview inventory which would surmount some of these limitations. The inventory was used in connection with an experience in teaching introductory multiplication, but it is felt that the basic approach has wide application. The purpose of the inventory was to present 3rd grade subjects with a number of multiplication examples for each to do, which, together with careful questioning, would reveal whether or not they demonstrated a rational, mechanical, or rote memory type of understanding of the multiplication operation.

Gray gives an explanation of his selection of items, developing the scoring techniques and criteria, the selection and training of the interviewers, and the strengths and limitations of the inventory. He concludes that this type of individual interview gives fairly sound evidence of varying levels of understanding that do exist among children; conventional tests of speed and accuracy do not provide this type of information necessary for evaluation of the outcomes of the new arithmetic program. (108)

Moody sets forth examples of items which might be used in a paper-and-pencil test to obtain evidence of pupils' thinking. If tests are to attempt to measure a person's ability to handle mathematical concepts, says Moody, we should select and develop questions which indicate how he "thinks" about mathematics, instead of what he recalls.

Example:

$$2^1 + 2^1 = 2^2$$
$$2^2 + 2^2 = 2^3$$
$$2^3 + 2^3 = 2\_$$
$$2^4 + 2^4 = \_\_$$
$$\_\_ + \_\_ = \_\_$$

(a.) Fill in the blanks above.

(b.) From part (a), a student might conclude that $2^n + 2^n = 2\_$

Moody further explains that it is more enlightening to study a person's concept, for example, of an angle as related to triangles and other geometric forms than to ask him to recall a possible series of meaningless words memorized to define such a figure. (109)

Ashlock, concerned about arithmetic achievement tests being computationally oriented, undertook to develop a test of understanding of selected properties of the number system in the primary grades. The test was so constructed so it was suitable for use with children in grades 1 and 2 and yet have high degrees of reliability and validity. Ashlock gives a full description of his purpose, procedures, findings, and conclusions. Some of his sample items might be worth our inspection. The orally presented directions had repetition built into them.

Ordinal Number Concepts

"Place your finger on D. Now in the row of numerals after the "D" put an "X" on the numeral which shows the number nearest to 44 when you count by ones. Put an "X" on the numeral which shows the number nearest to 44...

D	39	41	50	49

Place Value

Place an "X" on the numeral which tells how many ones the "2" in the box really stands for.

A	$\begin{array}{r} 147 \\ +\ 91 \\ \hline 238 \end{array}$	2	20	200	2000

Commutativity for Addition

Draw a line through the row of pictures that shows the same idea as is shown in the box.

D	$\boxed{3 + 5 = 5 + 3}$
	# + □ = △ + #
	○ + △ = # + ○
	# + △ = △ + #
	# + △ = # + △

Identity for Addition

Put an "X" on the numeral which could be used instead of the question mark to make the example true.

C	$\begin{array}{r} \triangle\ \square \\ +\ \ ? \\ \hline \triangle\ \square \end{array}$	0	1	2	4

(110)

Phillips gives several examples with which most of us are familiar at least for ideas of enrichment. As a test of understanding a multiplication algoririthm for middle grades she suggests putting the following example-with-a-hole-in-it on the board:

$$\begin{array}{r} 482 \\ \times 36 \\ \hline 2892 \\ \\ \hline 17352 \end{array}$$

Then tell pupils to imagine that they have forgotten all the multiplication facts, but that by using what they see here they can find the answer to 482 X 36. If they sense immediately how to do this (17352 -

2892) they understand the multiplication algorithm, whether or not they usually write the '0' in the second partial product, says Phillips.

Another suggested procedure consists of asking pupils to discuss two related examples, such as these, item by item:

```
                            3
   247                     80 } 83
   X83              247 ) 20501
   741                    19760
 1976                       741
 20501                      741
```

They should use some scheme (perhaps colored chalk) to point out related items. Then we should ask questions: "Why do we have '1976' in the multiplication example and '19760' in the division example?" "Do you have to write the '0's' in '20501'?" (111)

Of interest to us in the study of Grafft and Ruddell of cognitive outcomes of the SMSG mathematics program are the techniques they employed in an attempt to evaluate (1) the children's understanding of the mathematical principles recognized as part of the structure of math, and (2) the thought processes of the children. The researchers focussed on the operation of multiplication to secure three types of information. I wish to direct our attention to the first two which were:

1. A measure of multiplication computational ability, using conventional achievement tests, categorized accordingly:
   a. Whole numbers
   b. Common fractions
   c. Decimal fractions
2. A measure of understanding the structure of mathematics, using two types of tests
   a. A noncomputational, objective test measuring certain principles and properties of multiplication. Multiple-choice and true-false test items concerned the following principles:
      1.) Commutative property
      2.) Associative property
      3.) Distributive property of multiplication over addition
      4.) Identity property of one
      5.) Inverse relationship with division
      6.) Multiplication as repeated addition
      7.) Place value and the decimal system of notation

b. An individual interview inventory which presented children with the task of completing multiplication examples and was followed by careful questioning in order to reveal the following:

   1.) Ability to apply an algorithm
   2.) Ability to explain a multiplication operation
   3.) Ability to explain placement of partial products and why partial products must be added
   4.) Ability to explain decimal point placement
   5.) Knowledge of the inverse relationship of multiplication and division
   6.) Ability to multiply common fractions
   7.) Knowledge of place value and the decimal system of numeration
   8.) Knowledge of the idea of the distributive, associative, commutative, and identity property of multiplication
   9.) Ability to apply a variety of methods in the solution of the examples (112)

Sueltz, Boynton, and Sauble mention the role of computation and how it should be taught for understanding before bringing attention to its evaluation. For measuring understanding, say these writers, it is most fruitful to combine all of the usual procedures of evaluation, 1) use of written tests, 2) observation of daily work of the pupils, 3) interview of pupils during and after their work, and 4) self-evaluation by the pupils. Written tests are useful when they are prepared for a specific purpose. Illustrations are given to demonstrate the particular value of observations and interviews because they reveal when and how errors have been made. It is again emphasized that measurement and evaluation should be stages of the complete learning process, not ends in themselves. The authors indicate that if pupils really understand principles involved, they should be able to extend the principles into both (a) more complex socio-economic situations and (b) exercises with large or peculiar numbers. Pupils may be asked to state a problem similar to a type of problem that uses a particular process. (113)

The writers designate several kinds of understandings associated with computation and list questions and exercises which might help in evaluation under each one. I will reproduce only one such question under each understanding.

1. Understanding the usefulness of a process, as, for example, that addition is used for combining and grouping.
   Questions such as the following one are suitable for this group of understandings and may be used in paper-and-pencil tests followed by discussion:
   a.) John has 7 marbles, George has 15, and Pete has 12. How can you find how many all 3 boys have?
2. Understanding based upon the relation of one process to another, as, for example, the relation of addition to subtraction or multiplication.
   a.) Karen's allowance for 7 weeks was 25¢ a week. Find the total amount of her allowance in 2 different ways.
3. Understanding dealing with technical mathematical relationships in a process, such as "carrying" in addition
   a.) What happens when the 9 in the space in a speedometer changes to the next higher number? How is this like carrying in addition?

   This exercise and the others given serve best with interview and discussion to show degrees of understandings by individuals.
4. Understanding which depends upon a general mathematical sensing of number relations and leads to appreciation of reasonableness in results, as, when a number is multiplied by 4 the answer should be 4 times as large not 40 or $\frac{1}{4}$ as large.
   a.) Pick out statements that are always true. Make examples to illustrate those which are sometimes false.
      1.) When a group of numbers is added, the sum is always greater than the largest number.
      2.) The product of 2 numbers is more than the larger of the 2 numbers.
      3.) Any number divided by itself has a quotient of one.
      4.) When a number is multiplied by one, the product is one larger than the number.

5.) When the same number is added to both terms of a ratio, the value of the ratio is increased. The reverse is true when the number is subtracted. (114)

Sueltz et al. emphasize that observation, discussion, and interview serve better than paper-and-pencil tests in evaluating pupils' ability to understand the principles and procedures he uses in computation. Paper-and-pencil tests readily yield marks or scores but the writers point out that a teacher who works closely with his or her pupils should be able to form judgments of their understanding of their understanding of computation which is as fully reliable as scores obtained from written tests. A rich background in mathematics is required on the part of the teacher if understanding as well as mechanical skill in computation is to be evaluated.

The authors continue with a demonstration that tests in understanding may be constructed so they measure degrees of understanding and are thus useful for diagnostic purposes. The sample items presented in 13 categories suggest a.) types of items that can be used, and b.) the range of materials that can be covered. The items are constructed so they may be used in a paper-and-pencil test, but the teacher is urged to observe pupils while at work and interview afterward to discover why a pupil rejects some answers and why he finally selects what he considers correct. (115)

The writers advocate measuring a pupil's sensitivity to number in a social situation and even his habit of using numbers effectively in such situations. Formal tests reveal only a portion of a pupil's desire and ability to use his mathematics information in social and economic situations. The ideal way to judge a pupil's sensitivity is by observation of his behavior in his accustomed environment. It is difficult to observe children in their natural out of school habit, but school situations are available and ought to be utilized, such as the measurement, estimation, and judgment which goes into planning for a Valentine box. Written tests are particularly useful in measuring a pupil's abilities to use mathematics if tests are followed by discussion, and thus serve as astep in the teaching process as well as a measuring instrument. Sets of "multiple-choice" exercises are suggested under the classifications of:

1.) Knowledge of vocabulary of quantitative terms
2.) Knowledge of business practices and economic uses
3.) Interpretation of data presented in graphs and tables
4.) Usefulness of quantity in many aspects of life
5.) Using numbers in normal and vicarious experience: judgment
6.) Disposition to extend one's sensitiveness to the quantitative (116)

Earlier the different learning outcomes as distinguished by Ahman et al. were presented. They elaborate on each outcome and give suggestions for the evaluation of its different aspects. I will only mention techniques which vary from others already discussed.

Understanding processes - Have the child analyze each step of an operation with paper-and-pencil, using a code or set of abbreviations.

Example: Addition +; Subtract -; Multiply x; Divide ÷; Transform T; Remember R; Write in answer W

37			X	4 sevens, 28 ones
x4	x TRWX + W	or	T	2 tens, 8 ones
148			R	2 tens
			W	8 ones
			X	4 thirties, 12 tens
			+	12 and 2
			W	14

Pupils can help develop the code, providing further opportunity for the teacher to evaluate their understanding of the process.

Mental Arithmetic

Brighter children may workout shortcuts for working out algorisms.

Make use of the abacus. Since rods and beads merely tally results, a good understanding of the basic process is necessary in order for the pupil to manipulate the abacus.

Working with Fractions

Evaluative techniques using a number line, decimal ruler, or large decimal scale

Denominate numbers (measurement function)

In early school years, children tell what comparisons they make (big compared to what?)

Have pupil measure a desk, etc., explaining as or after he performs.

Pose questions and tape responses for objective evidence

Have plane figures in the classroom measured for their perimeter and area

Use paper-and-pencil tests for diagramming measurements to scale, blocking out a given area on squared paper, reproducing shapes of plane objects

Seeing relationships

Present a number table of appropriate material and ask the child to cite as many relationships as he can.

Have the child build magic squares which can be evaluated in terms of his ability to see relationships between numbers.

Ask the child to construct a table showing relationships he sees among a given group of facts.

Solving problems

Have the child explain a solution to a story problem without specified numbers or quantities, or merely have child write out the equation if numbers are included.

Have the pupil construct a picture or diagram of a problem solution.

Ask the child to make up a problem of his own. Careful analysis of the results shows what information the child selects from a situation, his ability to assign to think in terms of arithmetical processes.

Have the child develop a set of questions to ask himself when attacking a story problem.

Have the child restate a problem may provide insight into the thinking ability of the child.

Personal Use of Mathematics

Have the child give other illustrations of a given process, measure, or application from newspapers, magazines, radio, or television. Oral presentation or discussion by the individual also indicates real understanding.

History of Number

Present early-day methods of computation and ask the children to determine how the operation is performed, in order to evaluate his interest in and understanding of the real meaning of the process.

Attitude and Interest

To evaluate the degree to which a child enjoys estimating, playing with shortcuts, and number magic, observe how much he plays with such materials as cross number puzzles, "nimble" number tricks, riddles and stories with numbers, dominoes, "spinno," bean bags, hundreds board, counting block.

Keep a record of children who do enrichment work. (117)

Suggested under Brownell's fourth class of evaluative procedures are a variety of activities useful for teaching and evaluation.

Pupil reports, projects, and the like

Individual or group reports: "How Number Is Used in the ______ Bank", "How We Got Our Figure 7", "Lucky Numbers", "Arithmetic I Use on My Paper Route", "Where Our Measure Come From", etc.

Trips offer occasion for detecting awareness of the quantitative. Children who during note and after describe uses of arithmetic are clearly more advanced in their appreciation of the social significance of number than are others.

Events in the history of number can be dramatized, permitting evaluation of the level of quantitative thinking which has been attained.

In primary grades the ability to act out or picture events in verbal problems reveals understanding of the process meanings.

The preparation of scrapbooks, models, posters, and special exhibits, involving number and quantity may reveal the level at which children have achieved in their knowledge, understanding, and quantitative sensitiveness as well as teach them new concepts, skills, etc.

Meanings, understandings, and appreciation of arithmetic are most difficult to evaluate objectively, reliably, and validly by means of tests. Brownell stresses that these outcomes are more susceptible to evaluation, none too objective or reliable, but valid, when evaluation escapes the limits of testing and takes the form of one of the above procedures. The advantage of special reports, dramatization, and picturization is that they instigate behavior in which outcomes appear in natural relations and functional reality. (118)

Evaluating Attitude

In "Measuring Attitudes Toward Arithmetic" Dutton is concerned about the noticeable lacking of objective measures of attitudes toward elementary school subjects. In this article he reported the attitudes of prospective teachers toward arithmetic as determined by an objective evaluation instrument. The main findings reveal that attitudes toward arithmetic may be measured objectively and significant data may be obtained which will be helpful in the education of prospective teachers. Three of his conclusions which are most significant to us in considering evaluation are the following:

1. Techniques for measuring attitudes developed by Thurstone can be successfully applied to subjects taught in elementary school, perhaps laboriously but with desirable results.
2. The experience scale in the article gives significant information about attitudes toward arithmetic held by students, the intensity of the feeling, and general information necessary for guidance.
3. Feelings toward arithmetic are developed in all grades, but the most crucial are grades 3 through 6 and junior high. (119)

The Dutton Scale is used by Bassham et al. to investigate the relationship between pupil attitude toward arithmetic and pupil achievement in arithmetic, with individual differences in mental ability and reading comprehension held constant. In the sample of 159 sixth grade pupils, more than four times as many pupils with negative attitudes toward arithmetic were classified as 0.65 of a grade below expected achievement as were classified 0.65 of a grade above expected achievement. Almost 3 times as many pupils with positive attitudes overachieved 0.65 of a grade as underachieved that amount. (120)

Studies of Faust, Lindgren, and Shapiro are cited by Glennon as using varying means of measuring attitudes and finding positive relationships between attitudes toward arithmetic and achievement in arithmetic. (121)

Corcoran and Gibb treat the appraisal of attitudes in the learning of mathematics in chapter 7 of Evaluation in Mathematics. They examine the nature of attitudes and then methods of appraising attitudes. Self-report methods make use of questionnaires, rating scales, writing essays, and keeping diaries. Observational methods consist of the teacher observing and recording behavior that gives evidence of interests, attitudes, curiosity, and related reactions. A third approach, the interview, involves both the student and the observer. Attention is given to the interpretation of attitude data, its validity, reliability, and use. The importance of students' attitudes as indicators of what they have learned and as elements in motivation for further learning is pointed out. The writers state that in practice many judgments must be made on the basis of less than complete data. Care must be taken in interpretting any single piece of attitude evidence since it may provide an incomplete picture of a student's true position; different judgments should be checked against one another with the accumulation of items of evidence from various sources. (122)

Diagnosis

Diagnosis is considered by many of the writers previously cited in one way or another. We may gain insight into this problem by contemplating the writings of Glennon and Torgerson.

Glennon and Callahan bring to our notice that work in diagnosis in arithmetic has expanded concern from merely determining the kinds and frequency of errors in computional skills to growth in meanings and understandings basic to computional processes, growth in problem solving ability, growth in mental arithmetic, and growth in the ability to make quantitative judgments. They address to the concern in some of the newer thrusts in diagnosis, the complex relationship between growth in arithmetic development and affective factors such as anxiety, motivation, and attitude, mentioning studies of Bernstein and Ross and Brueckner.

With the realization of the complex nature of underachievent, states Glennon, methods of diagnosis must undergo change. (123)

Brueckner treats quite thoroughly the problem of diagnosis in arithmetic. Four general techniques for analyzing the nature of errors and faulty methods of work are discussed.

a.) Observation of pupil reactions
b.) Analysis of written work
c.) Analysis of oral responses
d.) The Interview (124)

A statement by Brueckner may give us insight into the scope of the ability required in diagnosing arithmetic learning.

> To diagnose arithmetic ability completely, the examiner must have a clear conception of the functions and objectives of arithmetic instruction, must be thoroughly acquainted with scientific studies of the factors that contributes to success in arithmetic, must know the symptons and causes of various unsatisfactory conditions, must be able to use effectively techniques for bringing to the surface facts concerning a pupil's disability and his thought processes which would ordinarily be unanalyzed, and must be able to interpret facts revealed by his study of the pupil and to suggest steps to correct the condition. (125)

Concerning diagnosis, Torgerson and Adams state that survey tests must be supplemented by 1.) observation, 2.) analysis of written work, 3.) diagnostic testing, and 4.) analysis of oral responses. They discuss diagnostic tests of computational skills and diagnostic tests of problem solving, examining specific tests within each group.

Studying the casual factors, Torgerson and Adams indicate that difficulties in arithmetic may be due to absense, inattention, insufficient practice, and other factors of a tempory nature or they may be due to a physical dificiency, insufficient mental maturity to understand the arithmetic process and techniques of problem-solving, poor work habits, or emotional problems. A frequent contributing factor is a distaste for arithmetic, accompanied by an attitude of defeatism. (126)

Glennon gives accounts of many different research studies to stress the part played by attitudes, anxiety, emotional disturbance, one's slef-concept, and other personality dimensions in causing certain difficulties in mathematics learning. (127)

An excerpt from Gertrude Hildreth's Learning the Three R's may bring us keener awareness of possible causes of difficulties which might otherwise escape our attention.

> The child may in spite of ample mental maturity fail in arithmetic if he has never learned to study or concentrate, if he can easily escape from exertion of effort, or if he is never held to any standards of competency...Emotional blocking resistance to instruction, lack of confidence, fear of failure, nervousness under speed pressure, distaste as a result of failure,

> and poor attention affect arithmetic just as seriously as other school activities. Impetuous, hasty, or nervously unstable children, those who are stubborn, recalcitrant, excessively painstaking, or nervously conscientious have emotional constitutions that predispose them to a deficiency in arithmetic unless they have exceptional mental ability. (128)

Using Evaluation to Meet Individual Needs

It has been previously indicated that for the evaluation process to be of worth it must be used to affect change. Throughout my presentation different follow-up procedures to evaluation and uses of the data obtained have been suggested in diagnosis, helping to plan for future teaching, reporting progress, etc.

It might be useful to give special attention to using the evaluation to help meet individual needs. Weaver described a teacher who did not have groups such that each had separate and independent instruction. Her groups were most flexible, being developed as the need arose, following her frequent exploratory work with the class as a whole. In her planning, the major emphasis was on differentiating in terms of the level or depth of learning rather than the rate of progress as we often conceive it. Interview data she gathered, which were described previously, were used to good advantage in reteaching. The data provided an idea of the strengths and weaknesses of each pupil and an indication of the level at which he might participate most successfully in the instructional activities. (129)

Bathurst advises that to help children to learn a particular skill the teacher must first make sure they are ready for the new learning. He must be assured that each youngster has already had some experience closely related to what he feels should be taught, or help him get such an experience. Secondly, the teacher selects a situation out of what the children are interested in doing. The situation should be real to the children, not something imagined or merely taken from a book. Then the teacher guides the children as they try the new learning. Other activities should be planned to extend and enrich the learning. We can see the necessity for evaluation preceding the first step and continuously throughout the process of instruction.

Bathurst relates the importance of the teacher giving children the opportunity to try ideas themselves and helping them achieve accuracy and feel success in whatever new phase of new learning they attempt. Evaluation is necessary in helping the teacher in his attempt to round out each lesson in such a way that pupils achieve their motive or goal. (130)

The title of Dumas's How to Meet Individual Differences in Teaching Arithmetic is self explanatory and indicative of a prerequisite of evaluation. Dumas points out the differences which exist in children in their knowledge of arithmetic and their ability to learn arithmetic. The teacher's problem as viewed by Dumas, is not how to dissolve differences, but how to teach so effectively that differences found will continue to extend themselves. To help avoid lockstep instruction, a plan is suggested to teach the same thing to all members of the class at the same time and with the same techniques. The plan involves quite flexible grouping, "inviting" children to join different groups as needs arise, allowing one group to work on items usually taught to children a grade or two below this group's present placement, and teaching another group new processes at a more rapid rate and at the same time giving them broad enrichment activities. Dumas gives further suggestions helpful in planning and carrying out group activities, allowing children help one another, utilizing incidental activities involving arithmetic and arithmetic possibilities in social studies and science units, and using games and puzzles. (131)

As we seriously contemplate the evaluation process and our role in evaluation it might be helpful to keep in mind these suggestions given by Brownell to teachers. The teacher is urged:

1. To know and understand outcomes set for instruction.
2. To know the various kinds of behavior which evidences growth toward these objectives and to train himself to detect this evidence.
3. To re-establish confidence in his ability to assess growth toward the more "intangible" outcomes.
4. To take advantage of the close relationship between teaching

and evaluating and seize every opportunity offered by everyday instruction to secure evidence of growth.

5. To realize that evaluation for some purposes is required rather seldom and may be managed by others than the individual teacher. (132)

EXPERIMENTING

WITH

CHILDREN

EXPERIMENTING WITH CHILDREN

Nature of the Project

In fulfillment of my desire to actually work with children in exploring the possibilities in mathematics evaluation, I was able to secure, with the aid of the superintendent of schools and the principal of one of the schools in Frankfort, Indiana, a classroom and other facilities and materials within the school, and four children with which to work.

My selection of children was narrowed by several early decisions. I wanted to work with children who might seemingly profit most from special attention and practice in mathematics. Since the time for our project was limited and I desired to spend much time working individually with each child, I was forced to limit the number of children involved and felt it desirable to have them near the same ability range and in the same grade. My final selection was made after a conference with a group of children suggested by the principal, Mr. Harry Moore, and two teachers. The four children chosen had completed the third grade, were of average intelligence, and each was having some difficulty in arithmetic, as evidenced by his scores on a standardized basic skills test taken the preceding year and by his grades in arithmetic.

After deciding upon the children I visited their parents to gain consent for their children's participation and gather information on the children and their home environment. All parents were enthusiastic about their children's participation and the benefit they might receive.

It was my intention to present the children with mathematical situations and problems in a variety of ways throughout the four weeks and observe their behavior in order to broaden my insight into mathematics evaluation. Throughout the four weeks from June 10 to July 5 I kept a daily record of all activities and a personal record for each child which included his accumulative record, his responses to the arithmetic reasoning test which I administered, anecdotal records of observations of his behavior in general and in mathematical situations, of his conversations, and of his explanations, and samples of his written work.

As a point of departure I administered the arithmetic reasoning section of the Iowa Tests of Basic Skills on our second day of meeting. This administration was most helpful in allowing me to observe each child in a stressful problem situation, and the responses provided information of possible weaknesses and strengths of each child. Items missed by a particular child were later presented to him in a slightly different manner in an attempt to understand his real difficulty. In this way I discovered that sometimes a child who has missed a particular problem has a better understanding of the situation and the process involved than another child who has happened to get the correct answer although his reasoning is faulty.

Each day my experimenting included working with or observing the group as a whole, and as individuals within a group, as they worked together in some mathematical activity such as a game of multiplication quizmo. Some time was devoted each day to presenting each individual with some kind of mathematical problem, talking with him, observing and recording his behavior and responses. Children not involved in individual work read books I had provided in the room, including some on number; played with games, many of which were of mathematical value; or amused themselves with construction paper, crayons, and clay. We also spent part of each day in outdoor or gymasium recreation, or other group activities such as singing, viewing filmstrips, or listening to some reading the children wanted to share.

It was often surprising that an evaluation experience came about in a most unexpected way at an unexpectedtime. Much was learned about each child which gave insight into his difficulty in arithmetic just by observing situations such as his physical recreation, conversations expressing his likes and dislikes, his agressiveness, his values, motivations, and by watching him engaged in mental activity in a game of "Pay the Cashier."

I recall one time when Dana was left to amuse herself with some clay. As I passed by her desk later on she pointed out twelve little globules of clay grouped into three's and exclaimed, "See, you can make sets out of the clay." This was a quite unexpected glimpse into her thinking on my part.

As you may already have been able to see, the great advantage I had in experimenting with procedures of evaluation with these children was their eagerness and willingness to help. They had been selected with as much understanding as I could give them as to what we would be doing throughout the four weeks and with the understanding that they would try to help me as much as possible. Activities were planned for the children such that they were as pleasurable as possible, and provided a feeling of accomplishment and success for each child. Needless to say, considering his background, each child was quite eager to receive individual attention.

The situation was unrealistic in that I had much more time than the classroom teacher to give to individuals, since I was mainly concerned with mathematics and was dealing with only four children. Thus you may say that such experimentation cannot be effectively used in the classroom. It has, however, in addition to the reading I have done, helped me gain keener insight into the thinking, understanding, attitudes, motivations, and behavior of children, how these can be assessed, and how they can be most helpful in our process of evaluation. I realize I have only acquired a little understanding, that I have only begun to study, but this experience has been a most enlightening, most inspiring and challenging beginning.

The complete record of the daily activities and the anecdotal records of each child would appear to be of little use to the reader. I have therefore reproduced only a few samples of the records I kept, hoping to illustrate some information or understanding which they provided. I have also included a list of specific procedures employed to gain information which might be used in evaluation.

Samples of Anecdotal Records

It is often possible to learn about a child's knowledge and his thinking merely by observing him respond to another child or by observing from a distance his behavior in a problem situation.

Wednesday, June 12, 1968

Pairs of children gave one another flashcards

Addition

Dondra - slow giving answer - figured in mind - got all right

Bill extremely slow - figured in mind, counted on fingers - missed 5 or 6

Kenneth explained - 5 + 7 means 5 + 5 + 2

Bill explained - 4 + 9 means 1 + 9 + 3
Bill explained - 8 + 5 = 7 + 6 = 13
Bill missed 8 + 1 several times - kept thinking it was times

Dana fairly fast - rote memory - missed 2

Kenneth fairly fast - missed 1

Subtraction

Dana - much difficulty
Dondra explained to Dana "What plus 6 equals 9?"
Dana didn't seem to be able to think in this way.

Kenneth - figured in his head
Asked him to explain his reasoning - said he just knew 12 - 8 = 4 - said he didn't think 4 + 2 = 6 to get 6 - 4 = 2 - said he just knew but it was hard to remember
Missed 5 - got the 2nd time (7-4, 4-1, 3-2, 3-1, 9-6)

Bill - very fast and accurate

Tuesday, June 11, 1969

Explained test was not for a grade and would not even be recorded - just for our own interest to see how well they can do

Allowed children to choose desks that fit - move their own desk from original position
Kenneth - turned desk to face chalkboard
Dana and Bill slanted theirs toward windows
Dondra - content with her original position

9:30 Children began test

Kenneth asked for 1st two to be explained - Hurried through pages - much distraction - looking around - marking answers hastily - finished test at 9:43 - made to check - finished - 9:49 - played with money game

Dana - appeared to be hurrying-finished at 9:47 - made to check - finished at 9:49 - read book

Dondra - had to have word given or question explained several times - counted on fingers - finished 9:59

Bill - looked around frequently - far away look - had not finished test at 10:00

It may be of interest to add that on this arithmetic reasoning test, Kenneth and Dana scored in the 7th percentile, Dondra scored in the 49th percentile, and Bill in the 72nd percentile, not finishing the last five items on the test.

It is often helpful in gaining understanding of a particular child just to ask questions and let the child respond freely.

June 27 - Dondra

What is arithmetic?	"Whole bunch of numbers, I guess. You have to do something with them, like add or subtract."
Do you like arithmetic? Why?	"Yes" "It's fun...to work problems"
Are there any you don't like?	"No"
Which do you like best?	"All of them"
Which are easiest?	"I don't know"
What is a number? a numeral? an odd number? an even number?	"Little marks" "A number" "Like 1,3,5,7 - Not even" "Like 2,4,6"
Does it always come after an odd number?	"Yes"
Is an odd number an even number plus one?	"Yes, and if you add two to an even number you get another even number."
Do you know what a fraction is?	"I've heard of it but I don't know what it is."
(Wrote 3, 2, $\frac{1}{2}$, 0.3 on the board) Which is a fraction?	She finally said "Two"
Gave her six sticks and asked her to give me half of them.	Without hesitation she handed me three.
Asked her why.	"Because six can be divided into three and three."
Who needs numbers? Why?	"Everybody" "For money, arithmetic, banking. You use numbers in almost anything you do."
How many numbers are there?	"From one to a thrillion."

Given manipulative aids, children can show in a concrete way their understanding of various concepts and processes.

Friday, June 21, 1968 - Bill

Provided cupcake containers of different colors.

Asked to show 5 + 3 - displayed 5 blue cups and 3 green ones.

Asked to show 5 - 3. Said he forgot how.

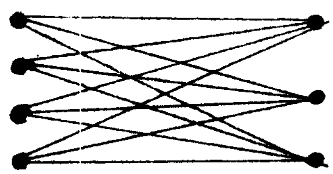
Asked how many to begin with - Answered 5

Asked to show subtracting 3 - Got out 3 more (Difficulty might have been in understanding term "subtract")

Asked to show 3 X 2 - Did immediately, explained he had 2 sets of 3.

Asked to show 8 ÷ 2 - Took out 8, divided them into 2 groups of four.

Presented him with picture of two columns of circles, one containing four circles, the other three. Had him draw one line from each of the circles in left column to each of the circles in right column and count the lines. Asked him if this illustrated some common operation he could think of. Didn't see the correlation until questioned as to how many in each column and how many lines there were. Then said it showed times.



4 3

12 lines

Monday, June 24, 1968 - Kenneth

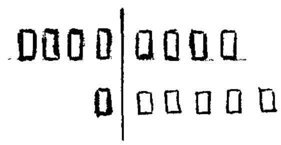
Provided with pocket chart divided in half, and strips of cardboard, red representing ones, blue representing tens, and green representing hundreds.

Asked to show 13 - easily put up correct strips

Asked to show how to add 8 more. Put on 8 more ones. Asked to write arithmetic sentences. After writing sentence he changed the

11 ones to one more ten and 1 one.

I placed strips on the chart.



Asked him to add these two numbers.

Put problem on the board. He explained well and demonstrated with the strips how he would change one of the tens to 10 ones in order to take away 5 ones.

Asked him to show with strips on chart how to divide 12 candy bars among 3 boys. Asked if he knew how many each would receive. "No." Divided by giving each one at a time. Wrote 12 ÷ 3 = on the chalkboard. Didn't know at first.

Gave him 4 strips. Asked to show me half. Asked to show half of 6 strips. Did both easily.

Anecdotal records of two children might be compared. These particular records seem to provide evidence of Bill's greater depth of understanding although neither boy was able to solve the problem.

Thursday, June 13, 1968

Kenneth - asked to read prices from a newspaper ad ranging from 77¢ to $3.75. Read correctly. Told he had $5.00 to spend and that he was to select as many items as he could buy with that amount. Kenneth wrote differences prices in several horizontal lines - added in his head - counted on his fingers -

25¢	10¢	5¢	5¢	3¢	
44¢	27¢				
44¢	77¢	$1.50	$2.50		etc.

Finally said he couldn't do it - showed little, if any, organization.

Bill - given same newspaper ad and directions. Immediately wrote down $5.00 - wrote small price underneath, erased it and changed it to $3.75. Much deliberation. Finally said he knew he would have to add other items to the $3.75 but he wasn't sure how to do it.

Thursday, June 27, 1968

Children were asked to make up a multiplication story problem. Evidences of their understanding may be thus revealed. Two examples follow.

Dondra

"I saw 2 things times 3 more things. How many do you have?"
2 X 3 = 6

Dana

"Judy and Bill went to the fair and Judy had 6 rides and Bill had

4 rides. How many were there?"   6 X 4 ____

24 rides

I asked Dana to reread her problem and asked her why she multiplied. "Because I had to make up a multiplication problem." Asked her if she shouldn't add. "Yes." Gave her 2 examples of simple multiplication problems and asked her to make up another problem.

"Janie and Jack went to their mother and they took her some candy in a basket. One piece of candy and one basket. How many were there?"

1 X 1 = ____

1 basket (What understanding does this reveal?)

Suggested Techniques

Flash Cards

Observe children responding to flash cards containing basic facts in addition, subtraction, multiplication, and division presented to them by classmates to help evaluate their knowledge, strengths and weaknesses in these operations. Listen to possible explanations given by the other child attempting to help the child figure out an unknown fact to gather evidence of the other child's understanding and thinking. Ask questions of the respondent to gain insight into his thinking and understanding of the operation.

Operation Quizmo

Each child has a card resembling a Bingo card. The number in the 25 squares may be answers to addition, subtraction, multiplication, or division problems depending upon the operation in which the game is being played. A problem, such as 3 X 5 is given and the children must solve the problem mentally to see if they have the number to be marked. By observing and listening to the children it is possible to gain knowledge such as who knows the facts by memory, who uses his fingers to figure, who counts in his head, who draws upon facts he does know and figures from there to get the answer. Often, the hints one child gives another in helping him figure out the problem himself is very enlightening.

The rule of "wordless playing" may be imposed to insure each child's thinking for himself.

Tell-Time Quizmo

Each child has a card resembling a Bingo card, but containing within each of its 25 squares a clockface showing a different hour. The caller chooses a card and reads the time given to the group. The child must be able to find the clock face showing the same time befoe he can mark it. Closely observing children playing the game shows who is having difficulty. The difficulty may be analyzed by having the child use a clock face with movable hands provided to show the time which has been read.

The child may be provided a movable-hands clock face when working with him individually. Have him read the hour shown on the clock as the hands are moved to different positions and have him place the hands in proper positions for a certain hour.

Cupcake cups, Pistachio nut shells, and Pieces of clay

Any small concrete objects have innumerable uses; the above are suggested because of their ease of storing. Have a child illustrate with such devices the process of a particular operation such as how he would group for addition or how he would have sets in multiplication. Ask the child to give you half, third, etc. of a given number of objects.

These same learning outcomes may be evaluated by having the child draw objects on paper or the chalkboard.

Place Value Chart

A pocket chart divided into thirds by pieces of yarn stretched vertically across chart. With the explanation that the 1st column represents ones, proceed to question the child. With strips of cardboard, colored differently depending on whether they represent ones, tens, or hundreds, have the child illustrate various numbers, explaining changes he makes when a number is added to the already existing illustration, etc.

Illustrate a number by placing strips in the pockets and have the child write the numeral on the board. Have the child demonstrate the changing of tens and ones, etc. as he explains different operations.

The chart and strips may also be utilized in having a child divide 12 "candy bars" among 3 boys, etc.

Tin Cans and Popsickle Sticks

The same procedures pertaining to place value which were employed with the pocket chart may be used with tin cans and popsickle sticks or tongue depressors. One can may hold the ones, another the tens, and so on.

These materials provide even more concrete experience as the child actually groups ten ones into a bundle to represent one ten, ten bundles to represent one hundred, and as he changes the bundles back to tens or ones.

Films

View a film of mathematical value. Discuss concepts suggested in the film to help obtain information about a child's understanding.

A filmstrip may serve as preparatory to another mathematics evaluation activity. For instance, a filmstrip on measurement might be followed by having children perform several activities with a ruler. A child who seems to have a good understanding of a certain aspect of measurement in a discussion following a filmstrip viewing may be found to have a little difficulty when actually confronted with a real problem situation.

Graph Paper

Present the child with a certain area of squares blocked off on a piece of graph paper.

Ask the child how he might find out how large the area is without a ruler. Ask questions to help him think about how many are on each side. Present larger areas so that counting the squares is impractical if the child does not show understanding of finding area. A child's explanations may reveal unexpected thought processes.

Flannel Board

Using three different groups of felt objects such as ducks, birds, and fish have the child make as many sets as he can with each set containing two different objects and no two sets being alike. Proceed to find evidence of the child's understanding of the process of multiplication.

These felt objects and flannel board may be used for children's manipulation to gain evidence of their understanding of the other operations or mathematical concepts.

Magic Squares, Crossnumber Puzzles and the Like

Observe child given squares or puzzles to complete to examine such things as knowledge of basic facts, mental arithmetic, and thought processes. Have child make up his own squares or puzzles. These also help in evaluation of interests.

Standard Test

Use items missed on a standard test as a basis for constructing other situations using more concrete objects to discover real understanding of the child,,his difficulties, and causes of his difficulties. To evaluate the real understanding of child regarding an item on measurement, have him draw how he thinks a cup, pint, and quart would look in relation to one another. Discuss his ideas.

Story Problems

Present child with a story problem, written or spoken. Have him explain how he would find the solution and why, and observe him as he solves the problem, whether mentally or with pencil and paper. If child has difficulty help him explain the problem himself by asking thought provoking questions about the problem situation.

Have child make up story problem. The problem may be designated to contain one particular operation, or to pertain to something the child saw on a walk around the block. The child also might be given ideas from various numbers and names of objects, and places listed on the board from which to choose, or from a picture presented or a well-known song. Observe the ease with which the child constructs the problem and the depth of understanding represented in the written problem. Problems may be given to other classmates to work.

Revealing information may be obtained by having the child draw a picture illustrating a particular problem.

Checking Problems

Ask child to check a problem using paper and pencil to obtain evidence of his understanding of the processes and their relation to one another. Questions may be asked the child having difficulty in knowing how to check a given process to obtain evidence of his understanding.

Newspaper Ads

Present the child with a newspaper ad containing many prices for various items. Have the child read the prices, and perhaps explain what the decimal point means or why a cent sign is used in place of a dollar sign.

Tell the child he has a certain amount of money and that he is to select as many items as he can possibly purchased with that amount.

Observe this procedure in determining which and how many items he can select. Have him explain his reasoning.

Geometric Figures

Have a child draw and construct from paper all the geometric figures he can think of. Ask him to explain and name these and others.

Present the child with several different wooden geometric forms one at a time, having him try to name them or express what he knows about them. The child might be asked to choose from a list on the chalkboard or a paper to label each model.

Puzzlers

Give the child a puzzler which involves his following directions to perform various operations in order to come out with a given number. This provides good opportunity to observe the child's ability to follow directions, his ability in performing certain processes, specific difficulties he may be having, and interest in number.

Number Line

Present the child with a number line. Have him show and explain how it may be used to help in the process of adding, subtracting, multiplying, and dividing, or in comparing fractions.

Interview

Besides interviews which may accompany presenting the child with problem situations, have "talks" with the child letting him respond freely to questions, such as:

"Who needs numbers? Why? What for?"

"How many numbers are there?"

"What is arithmetic?"

"What is a number?"

"What is an odd number?"

"What is a fraction?"

"Do you like arithmetic? Why, why not?"

"What kind of problem do you like?"

Response to a given question will often suggest further questions. Interviewing in such a way is very helpful in learning about a child's understanding, interest, attitudes, and motivations.

OUR CHALLENGE

TO

EFFECTIVELY EVALUATE

OUR CHALLENGE TO EFFECTIVELY EVALUATE

Reading in the area of evaluation, and particularly in mathematics evaluation, and working with children, I have deliberated over the problem of evaluation and the role I will play as a teacher. If my treatment of the subject seems perhaps too personal, it is a result of my feeling of great personal responsibility, which I am concerned that all persons involved in the education process possess.

Through the documentation provided by my reading, I have attempted to emphasize that as teachers are responsible for providing learning situations for the child, it is vital that teachers be concerned with the whole child and with the total development of each individual child. This brings the necessary condition of evaluating the whole child. For this reason mathematics evaluation cannot be viewed as an entirely separate entity as we might wish.

Evaluation and instruction of the whole child are vital to one another in order to have effective instruction or evaluation in any particular area. They may be viewed as components of a self-perpetuating cycle, each following and leading to the other. Thus we see that instruction and evaluation are continuous processes taking place simultaneously.

We should recognize not only that measurement, evaluation, and instruction are interrelated aspects of effective teaching, but that it is impossible to separate evaluation and teaching, and we should not try to do so. Torgerson and Adams were especially emphatic that evaluation should not be seen as an end activity to be done by some external agency, such as a teacher, ona specific calendar date. Evaluation is conceived as a necessary ingredient in all learning, to be engaged in by the learner and the teacher cooperatively, impossible of separation from other phases of the total learning process. (133)

The detrimental effects of evaluation when thought of as an end in itself are described in some of the readings. The teacher should take advantage of the close relationships between teaching and evaluation, seizing every opportunity offered by every day instruction to secure evidnece of growth, as Brownell suggests. (134)

As we survey a classroom of children we may be overwhelmed by the widespread individual differences in ability, achievement, attitude, and interest, to name only a few. These differences the teacher must seek to understand before he can make provisions for stimulating learning experiences for each child. In addition to the great variability within a class is the variability of strengths and weaknesses within an individual child.

Viewing the teacher's job as an evaluator, Torgerson and Adams distinquish two essential facets, each demanding constant attention. One facet has its orientation in the group instructional program and the extent to which the class as a whole is achieving its goals. The other has its orientation in the study of individual children, in the diagnosis of their growth lags and the discovery of important causal factors for such lags or deficienceies. Both of these approaches are significant aspects of evaluation and indispensable to good teaching. (135)

Torgerson and Adams explain the reasons for the teacher having the key role in measurement and evaluation today. Instructional and child study uses of measurement are recognized more clearly, and an increasing number of techniques and instruments are now available to the teacher. The teacher's greater responsibility and greater latitude in appraising the worthwhileness of those experiences. The teacher has access to background information about pupils which is necessary supplementary data, checking on hypotheses concerning causal factors, and trying out courses of action based upon such hypotheses. (136)

These I feel are important points to keep in mind. We should recognize that as greater knowledge and new techniques and instruments are put at our disposal and that as we are given greater responsibility in planning educational experiences, we also have a greater responsibility to evaluate more effectively, using all the knowledge and materials available to us.

Considering our responsibility to each child, we recognize the importance of finding methods of studying individual children such that evidence of specific causes of particular difficulties may be gained. Ten children may give incorrect solutions to the same item on a paper-and-pencil test as a result of ten different causal factors. It is necessary to understand the real cause of the difficulty before we can begin to effect a change.

The relationship between growth in arithmetic and affective factors such as anxiety, attitude, and motivation has been brought to our attention through the readings. The reader may refer to the causal factors contributing to various difficulties in arithmetic enumerated by Hildreth, given in this paper on page 58. I am concerned that when teachers evaluate a child in mathematics they pay little attention to such things as the child's never learning to concentrate, his possible emotional blocking, his lack of confidence, or his distaste for arithmetic as a result of failure. If we realize that a child is doing poorly in mathematics because he is nervously unstable, or perhaps because he is very hasty, we can begin to provide situations better suited to his particular needs, which may make all the difference in the world in his future development. This is one phase of evaluation in which I feel there is a general lack and a great need for deeper involvement and concern on the part of the teacher.

As implied earlier, studying and understanding children is essential to good teaching. As we discover a greater variety of differences and a wider range in degrees of understanding among individuals, our problem of meeting needs becomes increasingly greater. These differences discovered through evaluation should be utilized in planning more effective instruction. I find particularly challenging Dumas's statement that our problem is not how to dissolve differences, but how to teach so effectively that differences we find will continue to extend themselves. (137) Thinking in these terms, I believe we can begin to see evaluation in the proper perspective. Evaluation is a means to gain understanding of particular aspects of one's knowledge, understanding, thinking, particular abilities, achievement, attitudes, motivations, interests, and personality dimensions in order to provide the best possible learning situations for each individual. Mathematics evaluations should take into consideration all of these aspects, however unrelated they might seem at first, to create the most stimulating experiences for each child.

What we do to help meet individual differences once we have discovered them will depend upon circumstances, and upon our own ingenuity and industry. We should capitalize on a child's special abilities or interests by allowing him to use these to further his learning and increase and broaden his interests.

Information gathered from evaluation procedures may be the most valuable information in helping us plan our future teaching and re-teaching. Studying the children, we can teach to overcome some of their causes of difficulty, we can revise methods and try new ones for more effectiveness, we can gain better understanding of the children's levels of learning to know what might reasonably be expected of them. We can obtain such evidence of a child's understanding that we know where he can participate most effectively in the instructional process, when he can best contribute to a class discussion or project.

Evaluation, as Rothney has indicated, needs to show appreciation for the poorest pupils' good qualities while pointing out the weaknesses of the best pupils. (138) When evaluation can do this, we may plan ways in which pupils can be helped to overcome weaknesses and use their strengths more effectively.

Through the process of evaluation the teacher may increase the effectiveness of his teaching by restating goals of instruction as he sees them in his own teaching-learning situation. These goals or objectives should be stated in terms of changes in pupil behavior which the teacher strives to bring about. After each goal is stated in behavioral terms, the teacher should note situations in which he has the opportunity to observe and evaluate each type of pupil behavior. Only then is he prepared to select or devise means for collecting evidence about the behavior in question.

It is most important that we have well defined goals of instruction. These are not only the basis for our lesson and activity planning, but the very important learning outcomes we seek to evaluate. Therefore these objectives should be stated so they can be evaluated, and, according to Rothney (139), evaluations in turn should be of such character as to include the objectives as stated.

Many techniques of child study need to be understood, and their uses, limitations, and advantages recognized. I have attempted to bring to notice various procedures in the preceding chapters and hope that this may prove a useful reference.

Standard tests are important tools in aiding teachers in understanding, but their use needs to be supplemented by other techinques of evaluation.

Paper-and-pencil tests have value in readily yielding a score or grade which means something to the giver and should be understood by others. On the other hand, as several of the sources indicated, the teacher who works closely with the pupils should be able to form judgments of their understanding and their growth toward the various outcomes of learning set forth. These judgments should be as fully reliable as, if not more reliable than, the scores obtained from written tests.

We gather from the writing of authorities in evaluation that, in general, observation, discussion, and interview serve better than paper-and-pencil tests in evaluating pupils' ability to understand principles and procedures he uses in computation. These techniques are certainly most helpful in the discovery of certain attitudes, motivations, thought processes, interests, and possible causes of difficulty.

In selecting or constructing evaluative techniques, our main concern should be the characteristic we wish to evaluate and the qualities of each technique under consideration. In the past there has been increasing emphasis on the objectivity of scoring. Many of the writers previously cited have pointed out that this emphasis has often led to a superficiality in measuring in that goals for which achievement is measured most easily and objectively are often given greater emphasis than some more important goals of education. Deciding upon the outcome(s) we seek to evaluate and the behavior which will give evidence of growth toward the particular outcome(s), we must take care in selecting the most effective evaluative procedure to meet our needs in this evaluation. The procedure should achieve a high degree of objectivity without sacrificing the more important criterion of validity, as Torgerson and Adams have suggested. (140)

It is important to be guided in our evaluation program by the experiences we accept as most valuable in the teaching-learning situation. This would naturally have to be true if we were not trying to separate the evaluation process from other aspects of instruction. If we believe that it is important for children to work with problems in real life situations, they should not only be given opportunities to do their own experimenting and discovering of solutions in their initial and developmental learning experiences, but should often be evaluated in such real situations, rather than in some artificial testing situations.

Feeling that it is important for a child to feel success in a phase of new learning he attempts, we should plan evaluative experiences which also provide for some feeling of success or satisfaction of achievement.

Projects calling for problems in arithmetic and problems arising in natural situations provide concrete experience in quantitative thinking, fact-gathering, and computation. Such situations may bring a feeling of success to children as they solve the problems and provide valuable information in evaluation.

Certain desired outcomes of learning may themselves be used as aids in evaluation, or at least be furthered through their use in evaluation. Developing skill in self-appraisal is an example of an educational objective which may be used in such a way. Evaluation is reflected in the simple act of a pupil deciding for himself that something he has done is either good or bad. This tendency is often curbed by grades or marks which make a child dependent on a teacher's evaluation to the extent that one criterion of success becomes acceptability to the teacher. This does not appear to me as one of the outcomes of learning in which we want to foster growth.

The teacher can promote self-appraisal by commending effort and the product or by asking the pupil if his work could be improved. If the teacher feels the child is wrong in his appraisal, he should not try to impose his own evaluation but provide opportunities for the child's growth and development in the skill of self-appraisal. Self-appraisal requires daily practice; through continuous practice the skill is slowly developed. The teacher should record a pupil's oral efforts at evaluation and encourage the pupil to compare his present efforts with past work. It might prove beneficial to have some written evaluations; a child might write a letter to his parents or the teacher regarding his evaluation of himself. This practice may not only provide helpful information concerning a child's strengths and weaknesses but give insight into his depth of understanding and some of his attitudes and values.

Most relevent is a statement from Baron and Bernard:

> Wherever possible children should evaluate their own work, judging accomplishment and progress by charts kept by each child. As children come to you with different capacities and training, the competitive system of marking most usually practiced becomes unfair competition and often leads to dishonest practices. It leads to pressure where it is least helpful and very often allows the bright child to maintain

himself with no exertion. (141)

The dates of publication of the books to which I have referred reveal that for over thirty years increasing emphasis has been placed upon less formal methods of evaluation, such as observation and interview. After all these years it seems that we are still falling short of using these procedures to their fullest advantage. When it comes to reporting on a child's progress or making a change in instruction, teachers too often feel that they must rely upon completely objective data gathered from more formal sources of evaluation.

I feel very strongly that a teacher should use his imagination, his creativity, his own full capacities and special abilities, as well as those of his pupils, when he seeks to evaluate. I would advocate individualized evaluation in the sense not only that each child is evaluated according to his individual abilities and achievements but that each teacher use his own original ideas, as well as those gathered from other sources, and that he use varied approaches to evaluation with different children.

Two points of Brownell's important for the teacher to consider as he strives toward effective evaluation are that he should 1.) know the kinds of behavior which evidence growth toward specific objectives of instruction and train himself to detect this evidence, and 2.) re-establish confidence in his ability to assess growth toward the more intangible outcomes. (142)

The teacher should ever be gathering ideas for and information about new means of evaluation from his reading and his daily contact with the pupils. Much can be learned by keeping abreast of current research which is brought to our notice through such agencies as Riedesel and Pikaart's section on research in every issue of The Arithmetic Teacher, "Research on Mathematics Education..." for a particular year, produced yearly in the same periodical, or Glennon and Callahan's Guide to Current Research. It is also possible to construct most valuable evaluation procedures from ideas stimulated by articles suggesting specific instructional methods, which may or may not have to do with mathematics, or by pages within a mathematics enrichment workbook.

We should not be satisfied with the dull and commonplace, or the tried-and-true, in evaluation any more than we are satisfied with such methods in our teaching. Mathematics evaluation is the place for bold experimentation as far as varied techniques and materials are concerned. We have much to gain through constant experimentation with different means of evaluation. As we try our wildest ideas with children, we will frequently be surprised by information obtained, if not directly a result of our intended procedure, from avenues opened unto us by the children.

Evaluation is not, and cannot be, neutral. It is by its very nature and name, dependent on expressed values, and for this reason contains many subjective elements. Only when it is confused with a narrow interpretation of evaluation such as objective testing, can evaluation be said to be impartial. In this case, the real meaning of evaluation is lost and the way is open to undesirable situations in which the individuality of the child is forgotten. Let us not be so concerned about evaluating in such a way that other evaluators will obtain the same evaluation, or that each child receives the same kind of evaluation. Our chief concern should be that each evaluation aid in helping provide for the needs of each individual child and in our particular circumstances. With this end in view, let us accept the challenge to evaluate effectively ourselves, our educational program, and each individual child on the basis of his total personality.

FOOTNOTES

FOOTNOTES

1. Theodore L. Torgerson and George Sachs Adams. Measurement and Evaluation for the Elementary School Teacher. (New York: The Dryden Press, 1954) pp. 7-9.

2. Harold G. Shane and E. T. McSwain. Evaluation and the Elementary School Curriculum. (New York: Henry Holt and Co., 1958) pp. 400-402.

3. Jacob Wayne Wrightstone; Joseph Justman; and Irving Robbins. Evaluation in Modern Education. (New York: American Book Co., 1956) pp. 10-12.

4. J. Stanley Ahman and Marion D. Glock. Evaluating Pupil Growth. (Boston: Allyn and Bacon, Inc., 1963).

5. Denis Baron and Harold W. Bernard. Evaluation Techniques for Classroom Teachers. (New York: McGraw-Hill Book Co., Inc., 1958).

6. Wrightstone, op. cit., p. 3.

7. J. Stanley Ahman; Marion D. Glock; and Helen L. Wardeberg. Evaluating Elementary School Pupils. (Boston: Allyn and Bacon, Inc., 1960) pp. 7-9, 24.

8. Wrightstone, op. cit., pp. 3-4.

9. Torgerson, op. cit., p. 9.

10. Shane, op. cit., p. 59.

11. Ibid., p. 60-61.

12. Baron, op. cit., p. 239.

13. John W. M. Rothney. Evaluating and Reporting Pupil Growth. (Washington, D.C.: National Education Association, 1955) p. 3-4.

14. Shane, op. cit., pp. 61-62.

15. Ibid., pp. 62-66.

16. Wrightstone, op. cit., pp. 17-21.

17. Ibid., pp. 21-22.

18. Torgerson, op. cit., pp. 9-10.

19. Shane, op. cit., pp. 67-68.

20. Wrightstone, op. cit., pp. 38-41.

21. Ibid., pp. 41-42.

22. Harl R. Douglass and Herbert F. Spitzer. "The Importance of Teaching for Understanding" The Measurement of Understanding. 45th Yearbook. Part I. National Society for the Study of Education. (Chicago: University of Chicago Press, 1946) pp. 21-23.

23. Rothney, op. cit., pp. 30-31.

24. Baron, op. cit., pp. 224-227.

25. Wrightstone, op. cit., pp. 42-56.

26. Torgerson, op. cit., pp. 46-71.

27. Ahman and Glock, op. cit., p. 39.

28. Wrightstone, op. cit., pp. 42-56.

29. Torgerson, op. cit., 57-60.

30. Ahman; Glock; and Wardeberg, op. cit., pp. 45-46.

31. Baron, op. cit.

32. Torgerson, op. cit.

33. Shane, op. cit.

34. Wrightstone, op. cit.

35. Rothney, op. cic., p. 7.

36. Wrightstone, op. cit., p. 4.

37. Oscar Krisen Buros. The Third Mental Measurements Yearbook. (New Brunswick: Rutgers University Press, 1949).

38. Douglass, op. cit., pp. 19-21.

39. Torgerson, op. cit., pp. 40-43.

40. Baron, op. cit., pp. 256-258.

41. Ibid., pp. 229-236.

42. Douglass, op. cit., pp. 7-19.

43. Warren G. Findley and Douglas E. Scates. "Obtaining Evidence of Understanding." The Measurement of Understanding. 45th Yearbook Part I. National Society for the Study of Education. (Chicago: University of Chicago Press, 1946) pp. 45-64.

44. William A. Brownell and Verner M. Sims. "The Nature of Understanding." The Measurement of Understanding. 45th Yearbook. Part I. National Society for the Study of Education. (Chicago: University of Chicago Press, 1946) pp. 27-43.

45. Wrightstone, op. cit., pp. 355-375.

46. Baron, op. cit., p. 177.

47. Ibid., pp. 163-180.

48. Wrightstone, op. cit., pp. 380-391.

49. Torgerson, op. cit., p. 39.

50. Ibid., pp. 197-218.

51. Ibid., pp. 17-23.

52. Ibid., pp. 10-17.

53. Effie G. Bathurst. How Children Use Arithmetic. (Washington, D.C.: U. S. Department of Health, Education and Welfare, 1953) pp. 2-5.

54. Guy T. Buswell. "Methods of Studying Pupils Thinking in Arithmetic" Improving Mathematics Programs. ed. by M. Vere DeVault (Columbus, Ohio: Charles E. Merrill Books, Inc., 1961) pp. 298-306.

55. J. Fred Weaver. "Big Dividends from Little Interviews" Improving Mathematics Program. ed. by M. Vere DeVault. (Columbus, Ohio: Charles E. Merrill Books, Inc., 1961) pp. 322-348.

56. G. H. Miller. "How Effective Is the Meanings Method?" Improving Mathematics Programs. ed. by M. Vere DeVault (Columbus, Ohio: Charles E. Merrill Books, Inc., 1961) pp. 342-348.

57. Vincent J. Glennon. "Testing Meanings in Arithmetic" Improving Mathematics Programs. ed. by M. Vere DeVault (Columbus, Ohio: Charles E. Merrill Books, Inc., 1961) pp. 307-318.

58. Ahman; Glock; and Wardeberg, op. cit., pp. 278-279.

59. Shane, op. cit., p. 231.

60. Roland F. Gray. "An Approach to Evaluating Arithmetic Understanding" The Arithmetic Teacher. (13:187-191, March 1966).

61. William B. Moody. "Teacher Competence: What to Evaluate and How" The Arithmetic Teacher. (13:30-31, January 1966).

62. Robert B. Ashlock. "A Test for Understandings for the Primary Grades" The Arithmetic Teacher. (15:438-441, May 1968).

63. Ben A. Sueltz. "The Role of Evaluation in the Classroom" Evaluation in Mathmetics. 26th Yearbook. National Council of Teachers of Mathematics (Washington, D. C. 1961) p. 11.

64. Foster E. Grossnickle and Leo J. Breuckner. Discovering Meanings in Elementary School Mathematics. (New York: Holt, Rhinehart, and Winston, 1963) pp. 371-372.

65. Ibid., pp. 372-375.

66. Maurice L. Hartung. "Basic Principles of Evaluation" Evaluation in Mathematics. 26th Yearbook. National Council of Teachers of Mathematics (Washington, D.C., 1961) pp. 22-23.

67. Ibid., pp. 23-25.

68. Ibid., pp. 40-41.

69. William A. Brownell. "The Evaluation of Learning in Arithmetic" Arithmetic in General Education. 16th Yearbook. National Council of Teachers of Mathematics (New York: Teachers College, Columbia University, 1941) pp. 225-228.

70. Ibid., pp. 229-230.

71. Ibid., p. 230.

72. John L. Marks; C. Richard Purdy; and Lucien B. Kinney. Teaching Arithmetic for Understanding. (New York: McGraw-Hill Book Co., Inc., 1958) pp. 345-369.

73. Grossnickle, op. cit., 371.

74. Brownell, op. cit., pp. 232-235.

75. Marks, op. cit., p. 345.

76. Sueltz, op. cit., pp. 7, 8, 9, 14, 15, 20.

77. Hartung, op. cit., p. 40.

78. Grossnickle, op. cit., p. 378.

79. Vincent J. Glennon and Leroy G. Callahan. Elementary School Mathematics: A Guide to Current Research. Association for Supervision and Curricular Development, National Education Association (Washington D. C., 1968) pp. 83-84.

80. Ahman; Glock; and Wardeberg, op. cit., pp. 279-293.

81. Shane, op. cit., pp. 231-232, 241-246.

82. Max A. Sobel and Donovan A. Johnson. "Analysis of Illustrative Test Items" Evaluation in Mathematics 26th Yearbook. National Council of Teachers of Mathematics. (Washington, D.C., 1961) pp. 71-92.

83. Shane, op. cit.

84. Brownell, op. cit., pp. 231-232.

85. Ashlock, op. cit.

86. Glennon, op. cit.

87. Sueltz, op. cit.

88. Brownell, op. cit.

89. Ben A. Sueltz; Holmes Boyneton; and Irene Sauble. "The Measurement of Understanding in Elementary School Mathematics" The Measurement of Understanding 45th Yearbook. Part I. National Society for the Study of Education. (Chicago: University of Chicago Press, 1946) pp. 138-156.

90. Moody, op. cit.

91. Buswell, op. cit.

92. Weaver, op. cit.

93. William D. Grafft and Arden K. Ruddell. "Cognitive Outcomes of the SMSG Mathematics Program in Grades 4, 5, & 6" The Arithmetic Teacher. 15:161-165.

94. Mary Corcoran and E. Geraldine Gibb. "Appraising Attitudes in Learning Mathematics" Evaluation in Mathematics 26th Yearbook. National Council of Teachers of Mathematics. (Washington, D.C., 1961) pp. 105-122.

95. Sueltz, op. cit.

96. Wilbur H. Dutton. "Measuring Attitudes Toward Arithmetic" Improving Mathematics Programs. ed. bb M. Vere DeVault. (Columbus, Ohio: Charles E. Merrill Books, Inc., 1961) pp. 333-341.

97. Grossnickle, op. cit., p. 375.

98. Ibid., pp. 376-377.

99. Enoch Dumas. How to Meet Individual Differences in Teaching Arithmetic. (San Francisco: Fearon Publishers, 1957) pp. 5-6.

100. Brownell, op. cit., p. 235.

101. Shane, op. cit., pp. 246-254.

102. Marks, op. cit., pp. 351-369.

103. Glennon, op. cit., pp. 309-310.

104. Buswell, op. cit., pp. 299-303.

105. Ibid., pp. 306-307.

106. Weaver, op. cit., pp. 321-326.

107. Gray, op. cit., p. 188.

108. Ibid., pp. 187-192.

109. Moody, op. cit., pp. 30-31.

110. Ashlock, op. cit., pp. 438-441.

111. Jo Phillips. "Teaching New Math" The Instructor. 77:33-39, February 1966.

112. Grafft, op. cit., pp. 161-164.

113. Sueltz; Boyneton; and Sauble, op. cit., pp. 141-143.

114. Ibid., pp. 143-145.

115. Ibid., pp. 145-151.

116. Ibid., pp. 152-155.

117. Ahman; Glock; and Wardeberg, op. cit., pp. 282-294.

118. Brownell, op. cit., pp. 291-292.

119. Dutton, op. cit., pp. 341-342.

120. Harrell Bassham; Michael Murphy; and Katherine Murphy. "Attitude and Achievement in Arithmetic" The Arithmetic Teacher. (11:66-72, February 1964).

121. Glennon and Callahan, op. cit.

122. Corcoran, op. cit., pp. 105-121.

123. Glennon and Callahan, op. cit., pp. 81-82.

124. Leo J. Brueckner. "Diagnosis in Arithmetic" Educational Diagnosis 45th Yearbook, National Society for the Study of Education. (Bloomington, Illinois: Public School Publishing Company, 1935) pp. 291-293.

125. Ibid., p. 269.

126. Torgerson, op. cit., pp. 317-328.

127. Glennon and Callahan, op. cit., pp. 50-57.

128. Gertrude Hildreth. Learning the Three R's. (Philadelphia: Educational Publishers, Inc., 1936) pp. 460-461.

129. Weaver, op. cit., pp. 327-328.

130. Bathurst, op. cit., p. 8.

131. Dumas, op. cit., pp. 5-9.

132. Brownell, op. cit., pp. 242-243.

133. Torgerson, op. cit., p. 18.

134. Brownell, op. cit., pp. 292-293.

135. Torgerson, op. cit., p. 19.

136. Ibid., p. 23.

137. Dumas, op. cit. p. 5.

138. Rothney, op. cit., pp. 30-31.

139. Ibid., p. 5.

140. Torgerson, op. cit., pp. 57-60.

141. Baron, op. cit., p. 282.

142. Brownell, op. cit., pp. 292-293.

SELECTED

BIBLIOGRAPHY

SELECTED BIBIOGRAPHY

Ahman, J. Stanley, and Glock, Marion D. Evaluating Pupil Growth. Boston: Allyn and Bacon, Inc., 1963.

________, ________, and Wardeberg, Helen L. Evaluating Elementary School Pupils. Boston: Allyn and Bacon, Inc., 1960.

Ashlock, Robert B. "A Test for Understandings for Primary Grades" The Arithmetic Teacher. 15:438-441. May 1968.

Baron, Denis, and Bernard, Harold W. Evaluation Techniques for Classroom Teachers. New York: McGraw-Hill Book Co., Inc., 1958.

Bassham, Harrell; Murphy, Michael; and Murphy, Ketherine. "Attitude and Achievement in Arithmetic" The Arithmetic Teacher. 11:66-72. February 1964.

Bathurst, Effie G. How Children Use Arithmetic. Washington, D.C.: U. S. Department of Health Education and Welfare. 1953.

Brownell, William A. "The Evaluation of Learning in Arithmetic" Arithmetic in General Education. Sixteenth Yearbook of the National Council of Teachers of Mathematics. New York: Teachers College, Columbia University. 1941. pp. 225-267.

________, and Sims, Verner M. "The Nature of Understanding" The Measurement of Understanding. Forty-fifth Yearbook. Part I. National Society for the Study of Education. Edited by Nelson B. Henry. Chicago: University of Chicago Press. 1946. pp. 27-43.

Brueckner, Leo J. "Diagnosis in Arithmetic" Educational Diagnosis. Thirty-fourth Yearbook. National Society for the Study of Education. Bloomington, Illinois: Public School Publishing Company. 1935. pp. 269-302.

Buros, Oscar Krisen. The Third Mental Measurments Yearbook. New Brunswick: Rutgers University Press, 1949.

Buswell, Guy T. "Methods of Studying Pupils Thinking in Arithmetic" Improving Mathematics Programs. Edited by M. Vere DeVault. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961. pp. 298-306.

Corcoran, Mary and Gibb, E. Geraldine. "Appraising Attitudes in the Learning of Mathematics" Evaluation in Mathematics. Twenty-sixth Yearbook of the National Council of Teachers of Mathematics. Washington, D. C., 1961. pp. 105-122.

Douglass, Harl R. and Spitzer, Herbert F. "The Importance of Teaching for Understanding" The Measurement of Understanding. Forty-fifth Yearbook. Part I. National Society for the Study of Education Edited by Nelson B. Henry. Chicago: University of Chicago Press, 1946. pp. 7-26.

Dumas, Enoch. How to Meet Individual Differences in Teaching Arithmetic. San Francisco: Fearon Publishers, 1957.

Dutton, Wilbur H. "Measuring Attitude Toward Arithmetic" Improving Mathematics Programs. Edited by M. Vere DeVault. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961. pp. 333-341.

Epstein, Marion G. "Testing in Math: Why? When? How?" The Arithmetic Teacher. 15:311-319. April 1968.

Findley, Warren G. and Scates, Douglas E. "Obtaining Evidence of Understanding" The Measurement of Understanding. Forty-fifth Yearbook. Part I. National Society for the Study of Education. Chicago: University of Chicago Press, 1946. pp. 44-64.

Glennon, Vincent J. "Testing Meanings in Arithmetic" Improving Mathematics Programs. Edited by M. Vere DeVault. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961. pp. 307-318.

________, "Evaluation in Arithmetic and Talented Students" Improving Mathematics Programs. Edited by M. Vere DeVault. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961. pp. 330-332.

________, and Callahan, Leroy G. Elementary School Mathematics: A guide to Current Research. Association for Supervision and Curricular Development, N. E. A. Washingt n, D. C., 1968.

Grafft, William D. and Ruddell, Arden K. "Cognitive Outcomes of the S. M. S. G. Mathematics Program Program in Grades 4, 5, and 6." The Arithmetic Teacher. 15:161-165. February 1968.

Gray, Roland F. "An Approach to Evaluating Arithmetic Understanding" The Arithmetic Teacher. 13:187-191. March 1966.

Grossnickle, Foster E. and Brueckner, Leo J. Discovering Meanings in Elementary School Mathematics. New York: Holt, Rhinehart, and Winston. 1963.

Hartung, Maurice L. "Basic Principles of Evaluation" Evaluation in Mathematics. Twenty-sixth Yearbook of the National Council of Teachers of Mathematics. Washington, D. C., 1961. pp. 21-42.

Howard, Charles F. and Dumas, Enoch. Basic Procedures in Teaching Arithmetic. Boston: D. C. Heath and Co., 1963.

Marks, John L.; Purdy, C. Richard; and Kinney, Lucien B. Teaching Arithmetic for Understanding. New York: McGraw-Hill Book Co., Inc., 1958.

Miller, G. H. "How Effective Is the Meaning Method?" Improving Mathematics Programs. Edited by M. Vere DeVault. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961. pp. 342-348.

Moody, William B. "Teacher Competence: What to Evaluate and How" The Arithmetic Teacher. 13:30-31. January 1966.

Phillips, Jo. "Teaching New Math" The Instructor. 77:33-39. February 1966.

Riedesel, C. Alan. "Every Teacher Is a Researcher" The Arithmetic Teacher. 15:355-357. April 1968.

________; Suydam, Marilyn N.; and Pikeart, Len. "Research on Mathematics Education, Grades K-8, for 1967" The Arithmetic Teacher. 15:531-546. October 1968.

Rothney, John W. M. Evaluating and Reporting Pupil Progress. Washington, D. C.: N. E. A. 1955.

Shane, Harold G. and McSwain, E. T. Evaluation and the Elementary School Curriculum. New York: Henry Holt and Co. 1958.

Sobel, Max A. and Johnson, Donovan A. "Analysis of Illustrative Test Items" Evaluation in Mathematics. Twenty-sixth Yearbook of the National Council of Teachers of Mathematics. Washington, D. C., 1961. pp. 71-92.

Sueltz, Ben A. "The Role of Evaluation in the Classroom" Evaluation in Mathematics. Twenty-sixth Yearbook of the National Council of Teachers of Mathematics. Washington, D. C., 1961. pp. 7-20.

________; Boyneton, Holmes; and Sauble, Irene. "The Measurement of Understanding in Elementary School Mathematics" The Measurement of Understanding. Forty-fifth Yearbook, Part I. National Society for the Study of Education. Edited by Nelson B. Henry. Chicago: University of Chicago Press. 1946. pp. 138-156.

Torgerson, Theodore L. and Adams, George Sachs. Measurement and Evaluation for the Elementary School Teacher. New York: The Dryden Press, 1954.

Weaver, J. Fred. "Big Dividends from Little Interviews" Improving Mathematics Programs. Edited by M. Vere DeVault. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961. pp. 322-329.

Wrightstone, Jacob Wayne; Justman, Joseph; and Robbins, Irving. Evaluation in Modern Education. New York: American Book Co., 1956.

Materials Used with Children

Adler, Irving. The Giant Golden Book of Mathematics Exploring the World of Numbers and Space. New York: Golden Press, 1960.

Bendick, Jeanne. How Much and How Many The Story of Weights and Measures. New York: McGraw-Hill Book Co., Inc., 1947.

________, and Levin, Marcia. Take A Number. New York: Whittlesey House. McGraw-Hill Book Co., Inc., 1961.

Crescimbeni, Joseph. Arithmetic Enrichment Activities for Elementary Schools. West Nyack, New York: Parker Publishing Co., Inc., 1965.

Iowa Tests of Basic Skills. Multi-Level Edition for Grades 3-9. Form 1. Boston: Houghton-Mifflin Co., 1955.

Leaf, Munro. Arithmetic Can Be Fun. Philadelphia: J. B. Lippincott and Co., 1949.

May, Lola. Elementary Mathematics Enrichment. New York: Harcourt, Brace, and World. 1966.

National Council of Teachers of Mathematics. Enrichment Mathematics for the Grades. Twenty-seventh Yearbook. Washington, D.C., 1963.

Suppes, Patrick. Sets and Numbers Book 3A. New York: L. W. Singer Co., Inc., 1963.