DOCUMENT RESUME

ED 243 702 SE 044 473

AUTHOR	Speer, William R.
TITLE	Clinical Investigations in Mathematics Education. Thematic Addresses from the Fourth National Conference on Diagnostic and Prescriptive Mathematics.
INSTITUTION	Research Council for Diagnostic and Prescriptive Mathematics, Kent, OH.
PUB DATE	78
NOTE	107p.; Addresses delivered to the RCDPM National Conference (4th, College Park, MD, April 1-2, 1977). For related documents, see SE 044 464-465, and SE 044 479.
AVAILABLE FROM	Research Council for Diagnostic and Prescriptive Mathematics, 441 Beryl Drive, Kent, OH 44240.
PUB TYPE	Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Clinical Diagnosis; Computation; *Diagnostic Teaching; Educational Planning; Educational Research; Elementary Secondary Education; Learning Problems; *Literature Reviews; *Mathematics Education; *Remedial Mathematics; *Research Design; Research Needs; Speeches
IDENTIFIERS	Counting; *Mathematics Education Research; Research Reviews

ABSTRACT

Five invitational addresses presented at the Fourth National Conference of the Research Council for Diagnostic and Prescriptive Mathematics are included in this document. "The Nature of the Clinical Investigation" (Romberg and Uprichard) discusses the validity of clinical investigations and suggests guidelines for conducting such studies of instructional techniques. "A Clinical Model for Diagnosing Mathematical Deficiencies, (MD) Squared, Incorporating Educational Cognitive Style" (Speer) describes a process that takes into account factors that contribute to an individual's unique learning style and content deficiencies. "Feedback in Diagnostic Testing" (Engelhardt) reviews the literature on feedback as it relates to testing. Difficulties in existing research are identified, and needed research and changes in diagnostic procedures are suggested. "Counting Performance and Achievement: Some Preliminary Observations" (Callahan) considers the rote and rational counting performance of young children, with variables and relationships discussed. Finally, "A Perspective on the Future" (Heddens) briefly poses questions concerning the direction of research efforts and provides suggestions for attaining the goals of the Research Council. Three reaction papers are also included. (MNS)

CLINICAL INVESTIGATIONS IN MATHEMATICS EDUCATION

Thematic Addresses from
the Fourth National Conference on
Diagnostic and Prescriptive Mathematics

April 1-2, 1977
University of Maryland
College Park, Maryland

Editor:
William R. Speer
Publications Chairperson
Research Council for Diagnostic
and Prescriptive Mathematics
Bowling Green State University

Printed at Bowling Green State University, Bowling Green, Ohio 43403

Preface

In 1974, the First National Conference on Remedial Mathematics was held at Kent, Ohio, under the sponsorship of Kent State University and the Kent Educational Development Service Office. This conference, and two subsequent conferences in 1975 and 1976, provided impetus for the formation of the Research Council for Diagnostic and Prescriptive Mathematics--an organization created to promote stimulation, generation, coordination, and dissemination of research and development efforts in diagnostic and prescriptive techniques.

In 1977, the University of Maryland, in cooperation with the Maryland State Department of Education, hosted the Fourth National Conference on Diagnostic and Prescriptive Mathematics (a name change which reflects the aims of the RCDPM). This conference was planned for the purposes of (1) disseminating knowledge and techniques for diagnosing and prescribing activities for children and (2) providing a forum for critical discussion of research. Pursuant to these purposes, the conference included both general and research sessions encompassing talks, workshops, videotaped diagnostic sessions, sample analyses, and critical discussions on major problems in the field of diagnostic and prescriptive mathematics.

The Invitational Research Sessions, the subject of this report, included two types of major thematic addresses, namely, (1) reports of completed or ongoing research and (2) analyses of problems and needed research. Each thematic address was followed by a prepared reaction which, in turn, was followed by a panel discussion by selected speakers at the conference. The five thematic addresses are outlined in the following paragraph.

"The Nature of the Clinical Investigation" (Romberg & Uprichard) discusses the validity of clinical investigations and suggests guidelines for conducting such investigations in the study of instructional techniques and effects.

"$(MD)^2$--A Clinical Model for Diagnosing Mathematical Deficiencies Incorporating Educational Cognitive Style" (Speer) describes a diagnostic and prescriptive process that takes into account a variety of factors that contribute to an individual's unique learning style and content deficiencies.

"Feedback in Diagnostic Testing: Survey and Needed Research" (Engelhardt) reviews the literature on feedback as it relates to testing by questioning the "standardizing" procedure of prohibiting or delaying correct/incorrect feedback in individually administered diagnostic tests. Difficulties with existing research are identified and suggestions are made for needed research and changes in diagnostic procedures.

"Counting Ability and Achievement in Arithmetic" (Callahan) considers the rote and rational counting performance of young children. Variables associated with performance are discussed and relationships between counting performance and performance on various other arithmetic tasks are examined.

A final paper, entitled "A Perspective on the Future", given by James Heddens, RCDPM President, poses questions concerning the direction of research efforts and provides suggestions for attaining the goals of the Research Council.

The Research Council for Diagnostic and Prescriptive Mathematics is indebted to The University of Maryland, the Arithmetic Center at the University of Maryland, and the Maryland State Department of Education for their support. A particular note of congratulations is due to Dr. J. Dan Knifong for his outstanding efforts in organizing and overseeing the operation of the conference.

William R. Speer
July, 1977

Research Session Speakers

LEROY CALLAHAN, State University of New York at Buffalo, teaches courses on improvement of instruction in elementary mathematics. He has developed various protocols for diagnosing math learning difficulties, and directs Project CLIME, a clinic program in the Buffalo Public Schools.

LELON CAPPS, University of Kansas, acts as Senior Professor in Conducting Research on Error Patterns and Slow Learners in Mathematics. Dr. Capps is active in consulting with and creating diagnostic and prescriptive centers in public schools in Kansas.

JON ENGELHARDT, Arizona State University, has developed a graduate program at Arizona in diagnosis and remediation of mathematical learning difficulties, and has established a Mathematics Learning Center. Dr. Engelhardt is the Treasurer of the Research Council for Diagnostic and Prescriptive Mathematics.

JAMES HEDDENS, Kent State University, is author of Today's Mathematics. Dr. Heddens is responsible for the clinical program at Kent State and is President of the Research Council for Diagnostic and Prescriptive Mathematics.

LLOYD HUTCHINGS, Francis Marion College, in South Carolina, is best known for his development of low-stress algorithms for whole-number computation which have been featured in a recent National Council of Teachers of Mathematics yearbook.

TOM ROMBERG, University of Wisconsin, is one of the nation's leading researchers in elementary school mathematics. He directed the much-acclaimed DMP program which is based on research done at the Wisconsin Center for Research and Development.

WILLIAM SPEER, Bowling Green State University, is Director of the Mathematics Clinic and teaches elementary and secondary mathematics methods. Dr. Speer is currently serving as the Publications Chairperson for the Research Council for Diagnostic and Prescriptive Mathematics.

EDWARD UPRICHARD, University of South Florida, is Director of Graduate Studies in the College of Education. He teaches courses on diagnostic and prescriptive instruction in mathematics. Dr. Uprichard has published several articles, and made presentations on the treatment of learning difficulties in mathematics.

TABLE OF CONTENTS

The Nature of the Clinical Investigation

Thomas A. Romberg
University of Wisconsin-Madison

A. Edward Uprichard
University of South Florida

> For the experimentalist science progresses only in the laboratory; the theoretician views experiments rather as guides and tests for his models and theories; others see the most important task making counts and measures, or arriving at predictions, or formulating explanations; the fieldworker and clinician have still other viewpoints. All of them are right; what is wrong is only what they deny, not what they affirm. (Kaplan, 1964, p. 30)

We believe our audience to be investigators (primarily fieldworkers and clinicians) interested in diagnosis and remediation as they are related to the teaching of mathematics. We also assume that not all such investigators are aware of the appropriate methods of inquiry to study such problems. It is our experience that most educators have been trained to believe that to research is to experiment. Designing studies based on Campbell and Stanley (1963) and testing hypotheses using ever increasingly complex statistics have for many become the cannons of inquiry. Unfortunately this perspective of research is wedded to a particular philosophy of science—Baconian empericism—no longer considered the only method of inquiry in many philosophic circles. In particular, emperical inquiry in education has borrowed the tools and techniques from other disciplines—such as agriculture—which cannot capture the dynamic characteristics of learning children. At last year's conference two papers were presented which attacked the use of "standard" research methods in the study of instruction (Wilson, 1976 & Romberg, 1976).

In this paper it is our intent to go beyond those papers and examine the nature of one particular type of research as an alternative research methodology, the clinical investigation, and to offer some suggested guidelines for conducting clinical investigations focusing on diagnostic principles, and relating the results to instructional events.

In order to study the nature of the clinical investigation, it is necessary to briefly address the class of activities to which it belongs, namely, scientific inquiry.

Throughout recorded history man has attempted to obtain more knowledge about his environment and appraise its significance. In fact, within the very nature of man it is important that he make his world and his life meaningful and significant. To this end he has strived ceaselessly, but across the centuries there have been changes in the way in which man has obtained his information, organized it and analyzed it.

To understand the place of clinical investigations within the realm of scientific inquiry we believe that it is important that investigators realize that there are several sources of knowledge which man has used and still uses in obtaining knowledge. The first and most prevalent is personal experience. For example, a classroom teacher may find a particular teaching method useful with a certain group of students and may record that information for use in future years. There are many dangers in appealing to personal experience, generalizations can be drawn on insufficient evidence or too few examples, or incorrect conclusions can follow from prejudice or evidence being left out because it was not consonant with earlier experiences. Also there is always the danger of failing to recognize which were the salient features of a situation and which were irrelevant. However, we still base much information on personal experience.

A second source of information has and continues to be authority. In everyday life a certain amount of knowledge is handed down by parents, teachers, scientists, for example and must be taken on trust. Life is too short to test the validity of every view that is held. For example, it is commonplace within education to expect committees of scholars in a particular field to outline directions. This is an important part of how human beings gather information. However, just because it has been handed down and widely believed does not necessarily mean it is true. Generally speaking the expert is better informed in his field than other people because of his level of intelligence, training and experience, but experts disagree among themselves.

A third method of scholarly inquiry has been deductive reasoning. In ancient Greece the deductive method of reasoning was developed to a highly refined art. Euclid's elements remain as one of the pinnacles of human thought. Deductive reasoning is used in everyday life to solve a host of problems. The detective, the doctor, and the educational researcher use it. The diagnostician in particular uses it when he searches among evidence and symptoms and selects items which appeared at first to be unrelated and brings them together in a way that logically lead to a conclusion. Note, however, that deductive reasoning only enables us to deduce the consequences of what was already known. It does not yield any really new knowledge. If one of the premises is untrue the conclusion will be false. Thus, it cannot be relied upon exclusively.

The fourth method of inquiry is induction. In the seventeenth century Francis Bacon successfully attacked the Aristotelian method of sylogistic reasoning. His notion was that a scientist begins by carrying out experiments whose aim is to make carefully controlled and meticulously measured observations at some point on the frontier between our knowledge and ignorance. He systemically observes and records his findings and in the course of time he and other workers in the field accumulate a lot of shared and reliable data. As this grows general features begin to emerge and individuals start to formulate general hypotheses (statements of a law-like character) which fit all the known facts and explain how they are causally related to each other. The individual scientist then tries to confirm his hypothesis by finding evidence which will support it. If he succeeds in verifying it he has discovered another scientific law which will unlock more of the secrets of nature. Thus, the existing stock of scientific knowledge is added to in the frontier and our ignorance pushed back. The process begins again on a new frontier.

From the 17th century to the beginning of the 20th century Baconian empericism based on the fundamental notions of inductive inquiry were viewed as the method of scientific inquiry in most fields. In this century more and more philosophers of science when examining major scientific breakthroughs have found that the inductive method is not the method used by most scientists in forming generalizations and concepts about the world. In fact as Einstein stated "There is no inductive method which could lead to the fundamental concept of physics." Failure to understand this fact constituted the basic philosophic error of so many investigators in the 19th century. Dewey in 1933 proposed another method which he pretentiously titled The Scientific Method. His description combines both inductive and deductive reasoning in a sequential state of hypothesizing, logical consequences of hypothesis, more data, etc.

Now in the later part of the 20th century it has been posited that scientific inquiry does not lead to internal truth but rather to systematic doubt. Karl Popper (1963) and other falsificationists are arguing that the critical falsification is important and not critical confirmation. It is their notion that sciences are not bodies of established fact, rather they are practical statements about the level of knowledge in a particular area. This implies that theories are not bodies of impersonal facts but are products of the human mind. This makes theories personal achievements of an astonishing order. The implications of this stance are as follows:

1) How scientists arrive at a theory does not bear on its scientific or logical status.
2) Second, observation and experiments far from giving rise to the theory are partially derived from it and they are designed to falsify the theory.
3) At no point is induction a major part of the set of activities.

Instead, hunch, inspiration, illumination, and insight are as important as control. Explanation and prediction become critical aspects not verification.

Activities within scientific inquiry may be thought of in two contexts--the context of discovery or the context of justification. Researchers focusing on activities within the context of justification are guided by a "logic of proof." They are primarily interested in reasons for accepting or rejecting hypotheses (confirming or falsifying propositions). Researchers focusing on activities within the context of discovery are guided by a "logic of discovery." They are interested in possible reasons for generating or entertaining propositions. The results of such inquiry rather than learning facts or laws to follow are ideas to use, modify, and expand upon.

The clinical investigation belongs to a subclass of activities within the context of discovery. It is an inquiry process that is used to search for and examine patterns in the study of phenomenon(a) in order to generate an array of data and hypotheses rather than to validate conclusions. It is common for chance, intuition, and imagination to play important roles in searching for intended and/or unintended outcomes in this type of investigation.

Success with the clinical investigation depends on two aspects: Analytic/Synthetic Procedures and Methodological Techniques. Analytic/Synthetic research focuses on the development of guiding models, or paradigms, and explanatory theories (Wilson, 1976). These models and theories play key roles in designing the clinical investigation, collecting and interpreting data, and in formulating hypotheses.

Because education is of such universal concern in our society and everyone has been in schools, most everyone feels free and competent to speak about it in general terms. We are now overwhelmed with a literature of unsubstantiated general ideas and principles often based on firm prejudices and soggy arguments. Unfortunately, this literature is not new in education. Our history is an open record attesting to a lack of intellectual discipline. This is also true in studying diagnostic principles and trying to relate the results to instructional practice. But, we are not pessimistic about the state of affairs. Unsubstantiated claims are a natural part of inquiry. Our common problem is how to attack the belief structure where it is unsupported by data and systematic theory. The clinical study is well suited to this task for as Popper argues to falsify the proposition—all swans are white—we need to find only one "black" swan. No matter how many examples may be found to confirm the proposition only one counter example can falsify it. Thus, in the Popperian view, theory building begins with existing propositions derived from prior theories, hunches, biases, etc. Then through systematic attempts to falsify these claims, new patterns emerge.

Methodological techniques focus on the methods, measurement techniques, etc., used in scientific inquiry. That is, methodological studies are designed to describe and analyze methods, throw light on their limitations and resources, clarify their presuppositions and consequences, and relate their potentialities to the twilight zone at the frontiers of knowledge (Kaplan, 1964). Methodological research helps one to understand the scientific inquiry process itself rather than its products.

For what we are calling "clinical investigations" there are five distinct methods that investigators have used. They have been called structured individual interviews, Piaget's "method clinique," the teaching experiment, ethnographic case studies and process-development evaluation. All the methods are characterized by the investigator being clinical. That is, he focuses on a particular phenomena as he attempts to capture essential features of the event. This implies the investigator cannot be simply a "data cruncher." Instead he must be intimately involved with the event by observing it and quite often structuring it by the types of questions he asks.

The structured individual interview has long been used in studying mathematical knowledge. Weaver's (1976) selected reference[1] on the topic began in 1891 and continues today. Bechtel's study (1976) is a good example of this kind of investigation. He interviewed second graders by presenting them a structured set of physical division problems. He limited his questions in a systematic way and videotaped all the interviews. All interviews proceeded in the same way. The behaviors of all the subjects were then coded into categories and relationships examined.

Piaget's "method clinique" starts the same as the structured interview--with a carefully designed specific problem. It differs however in how the interview is carried out. In this technique the investigator is like a Freudian psychiatrist who probes the responses each subject makes. No two interviews are expected to proceed in the same way. The behaviors are then analyzed and patterns examined.

Like the two previous clinical techniques the teaching experiment starts with carefully designed problems. However, in this case rather than trying to discern pupil's knowledge or how they think the investigator is interested in how pupils learn or acquire new knowledge. Each subject is his own control in that his later performance is gauged in terms of earlier performance. Although not unknown in this country the "teaching experiment" has been a primary research tool of Soviet researchers. For example, Krutetskii's recently translated work on mathematical abilities summarizes voluminous information derived from teaching experiments (1976).

The ethnographic case study in contrast to the other techniques does not start with investigator control over the problem situation. Rather it is an attempt to discern the structure of naturalistic events. The methodology has been borrowed from anthropology--after all the anthropologist cannot restructure the culture he is studying. An example of this type of research may be found in the current Beginning Teacher Evaluation Study being carried out by the Far West Laboratory. Tikienoff, Berliner and Rest (1975) gathered ethnographic data on forty elementary classrooms while reading and mathematics were being taught.

[1] This set of references is in appendix A.

<u>Process-development evaluation</u> combines some features of the leading experiment, ethnographic studies with more standard assessment procedures in a time-series design (See Romberg & Fox, 1976; and Fox, et al., 1976).

Researchers engaging in clinical investigations must be able to (1) articulate the relationship(s) between their research and its theoretical base(s) and (2) justify the methods or techniques employed.

The Clinical Investigation and Instruction

Research on instruction could benefit by using clinical methods particularly teaching experiments and ethnographic case studies.[2] Research relating to instruction focuses primarily on information processing variables of individuals (or groups), personality variables, and learner-environment interactions. The nature of these variables is such that in many instances they are better studied using idiographic techniques in a clinical setting rather than under experimental conditions. Cronbach, realizing the difficulties of studying variables related to instruction, suggests the following approach:

> Instead of making generalizations the ruling consideration in our research, I suggest that we reverse our priorities. An observer collecting data in one particular situation is in a position to appraise a practice or proposition in that setting, observing effects in context. In trying to describe and account for what happened, he will give attention to whatever variables were controlled, but he will give equally careful attention to uncontrolled conditions, to personal characteristics, and to events that occurred during treatment and measurement. As he goes from situation to situation his first task is to describe and interpret the effect anew in each locale, perhaps taking into account factors that were unique to that locale or series of events... The special task of the social scientist in each generation is to pin down the contemporary facts (Cronbach, 1975).

Clinical methods fit the approach described above and thus lend themselves to the study of instruction.

The instructional event is an interactive process and is not static. The variety of participation responses often changes the content and direction of the event, some outcomes are often embedded within the process itself, each event is occurring within a larger and complex learning milieu, and the apparent content of the event shifts and changes when viewed from different personal and time perspectives. Presumably,

---

[1]Research on diagnostic principles could also benefit by using clinical methods particularly structured interviews or by using Piaget's "method clinique." Romberg (1977) has dealt with this issue in his other paper at this conference.

to understand what happens in an instructional event, one must have a description of the event itself which illuminates these characteristics. In particular, any event must be considered as a mixture and interaction of physical, social and moral-psychological environments. The physical space (time, place and objects) in which an event occurs is suggestive, facilitating and constraining, with respect to what occurs. What people do and say to and with others both 'direct' observable behavior (reactions to planned actions and questions) and 'natural' observable behavior (naturally occurring 'free' actions and responses) are part of every event. And finally, the intentions of all the participants shape and give meaning for each individual to each event.

In studying instruction, the researcher's task is to discern the patterns of interactions in order to describe the course of events and predict effects. He uses theoretical constructs of social interactions, or operational categories based on an empirical examination of the event to identify patterns. Also, effects are not just terminal outcomes, they include how ideas develop and change over time. Thus, the central problem faced by researchers interested in studying instructional events is to attempt to capture the patterns of interaction which actually occur.

Summary

If one's knowledge about diagnosis and instruction in mathematics is to improve, then we collectively must be more imaginative in our approach to inquiry. Our intent has been to stimulate you to consider an alternative method of inquiry--the clinical investigation--as a way of discovering practical propositions about what we know.

Some Suggested Guidelines for Conducting Clinical Investigations

General

1. Know your problem--"no one goes in cold."
2. Examine how others have carried out similar studies.
3. Plan to pilot ideas several times--be open to change.
4. Talk over ideas, plans, results, objectives...with others.
5. Assume that you will learn--do better next time.
6. Clearly define the role of the investigator.
7.
8.
9.

Instrumentation

1. Determine what type of instrumentation (anecdotal profiles, rating scales, structured interview protocols, questionnaires, check lists, observational systems, diagnostic tasks or probes, tests, etc.) best fits the nature of your problems and your objectives.

2. Instrumentation should clearly reflect the constructs or theoretical rationale related to your problem and your objectives.
3. Consider validity (content, concurrent and/or construct) of instrumentation.
4. Consider reliability of instrumentation.
5.
6.
7.

Observations

1. Clearly identify the dimension(s) of behavior or process to be observed.
2. Clearly identify subjects to be observed.
3. Clearly identify the setting in which the observations are to occur.
4. Carefully plan and by systematic in making 'direct' and 'natural' observations.
5. Train observer(s) on what and how to observe.
6. Consider using video tapes and recorders.
7. Consider reliability of observational procedure through inter-observer correlations.
8. Consider observing behavior(s) or process more than once--Time Sampling.
9. "live" observations should be recorded immediately or as soon as possible.
10. Consider making provisions for recording unusual or unexpected behavior during observed event.
11.
12.
13.

Data Analysis

1. Categorize data so that it will reflect your problem and purpose.
2. Make sure you have a sufficient number of data categories.
3. Make sure your categories are independent.
4. Make use of tables, frequency distributions, graphs, etc. when possible.
5. Make use of individual or group profiles if appropriate.
6. Make use of appropriate statistics.
7.
8.

Writing Results

1. Clearly state results--be specific.
2. Clearly relate results back to your problem and theoretical rationale. (How does your study advance the frontier of knowledge?)
3. List hypotheses generated by study for future research (if appropriate).
4. Identify unexpected patterns of behavior exhibited by individuals and/or groups.
5. Be careful--don't over generalize results of clinical investigations.
6.
7.
8.

REFERENCES

Bechtel, R. D. & Weaver, J. F. A preliminary investigation of the ability of second-grade children to solve division problems presented in a physical-object mode and in different partitive-measurement sequences. (Working Paper No. 194.) Wisconsin Research and Development Center for Cognitive Learning, 1976.

Campbell, D. & Stanley, J. Experimental and quasi-experimental designs for research in teaching. In N. Gage (ed.) Handbook of research on teaching. Chicago: Rand McNally & Co., 1963.

Cronbach, L. J. Beyond the two disciplines of scientific psychology, The American Psychologist, 1975.

Fox, G., Tabachnick, B., Romberg, T., Wehlage, G., Popkewitz, T., & Grant, C. CMTI impact study interim report. Madison, Wisconsin: Teacher Corps Evaluation Project, 1975.

Kaplan, A. The conduct of inquiry, methodology for behavioral science. Chandler Publishing Co., 1964.

Kruteskii, V. A. The psychology of mathematical abilities in school-children. Chicago: The University of Chicago Press, 1976.

Popper, K. R. Conjectures and refutations: The growth of scientific knowledge. New York: Basic Books, 1963.

Romberg, T. A. The application of "process evaluation" to corrective instruction: A suggestion. In J. Heddens & F. D. Aquila (eds.), Proceedings of the third national conference on remedial mathematics. Ohio: Kent State University, 1976.

Romberg, T. A. Second level diagnosis: Assessing underlying processes, not just objectives. Paper presented at the 4th National Conference on Diagnostic and Prescriptive Mathematics. University of Maryland, 1977.

Romberg, T. A. & Fox, G. Problems in analyzing dynamic events in teacher education. Paper presented at American Research Association Annual Meeting, April 1976.

Tikunoff, W. J., Berliner, D. C. & Rest, R. C. An ethnographic study of forty classrooms of the BTES known sample. (Technical Report No. 75-10-5.) San Francisco: Far West Laboratory, 1975.

Wilson, J. W. An epistemologically based system for classifying research and the role of clinical intervention research within that system. In J. W. Heddens & F. D. Aquila (Eds.), Proceedings of the third national conference on remedial mathematics. Ohio: Kent State University, 1976.

Weaver, J. F. The use of "individual interviews" in assessing pupils' mathematical knowledge or learning: A chronology of selected illustrative references. University of Wisconsin-Madison, April 18, 1973; revised 1976.

Appendix A

Originally prepared by J. Fred Weaver
April, 1973

Revised
April, 1976

The Use of "Individual Interviews" in Assessing Pupils' Mathematical Knowledge or Learning: A Chronology of Selected Illustrative References

1. Hall, G. Stanley, Contents of children's minds on entering school. Pedagogical Seminary 1: 139-173; 1891.

2. Gard, Willis L. A preliminary study of the psychology of reasoning. American Journal of Psychology 18: 490-504; October 1907.

3. Judd, Charles Hubbard. Studies of number consciousness. Psychological Bulletin 6: 42-43; February 15, 1909.

4. Buswell, Guy Thomas, & Judd, Charles Hubbard. Summary of educational investigations relating to arithmetic. Supplementary Educational Monographs, No. 27. Chicago: University of Chicago, June 1925.

5. Buswell, G. T. (with the co-operation of Lenore John). Diagnostic studies in arithmetic. Supplementary Educational Monographs, No. 30. Chicago: University of Chicago, July 1926.

6. Brownell, William A. The development of children's number ideas in the primary grades. Supplementary Educational Monographs, No. 35. Chicago: University of Chicago, August 1928.

7. Brownell, W. A. The techniques of research employed in arithmetic. In Report of the Society's Committee on Arithmetic. Twenty-ninth Yearbook of the National Society for the Study of Education. Bloomington, IL: Pulbic School Publishing Co., 1930. (Pp. 415-443)

8. Buckingham, B. R., & MacLatchy, Josephine. The number abilities of children when they enter grade one. In Report of the Society's Committee on Arithmetic. Twenty-ninth Yearbook of the National Society for the Study of Education. Bloomington, IL: Public School Publishing Co., 1930. (Pp. 473-524)

9. Buswell, G. T. A critical survey of previous research in arithmetic. In Report of the Society's Committee on Arithmetic. Twenty-ninth Yearbook of the National Society for the Study of Education. Bloomington, IL: Public School Publishing Co., 1930. (Pp. 445-470)

10. Buswell, G. T., & John, Lenore. The vocabulary of arithmetic. Supplementary Educational Monographs, No. 38. Chicago: University of Chicago, January 1931.

11. Brownell, William A., & Chazal, Charlotte B. The effects of premature drill in third-grade arithmetic. Journal of Educational Research 29: 17-28, Supplement 1-4; September, November 1935.

12. Brownell, William A., & Watson, Brantley. The comparative worth of two diagnostic techniques in arithmetic. Journal of Educational Research 29: 664-676; May 1936.

(Continued)

13. Brownell, William A. (with K. G. Kuehner & W. C. Rein). Learning as reorganization: An experimental study in third-grade arithmetic. Duke University Research Studies in Education, No. 3. Durham, NC: Duke University Press, 1939.

14. Brownell, William A. The evaluation of learning in arithmetic. In Arithmetic in general education. Sixteenth Yearbook of the National Council of Teachers of Mathematics. New York: Bureau of Publications, Teachers College, Columbia University, 1941. (Pp. 225-267)

15. Brownell, William A. (with R. A. Doty & W. C. Rein). Arithmetic in grades I and II: A critical summary of new and previously reported research. Duke University Research Studies in Education, No. 6. Durham, NC: Duke University Press, 1941.

16. Brownell, William A., & Carper, Doris V. Learning the multiplication combinations. Duke University Research Studies in Education, No. 7. Durham, NC: Duke University Press, 1943.

17. Brownell, William A., & Moser, Harold E. (with D. M. Frazer, M. F. Karr, J. H. Ostwalt, & J. F. Weaver). Meaningful vs. mechanical learning: A study in grade III subtraction. Duke University Research Studies in Education, No. 8. Durham, NC: Duke University Press, 1949.

18. Buswell, G. T. Methods of studying pupils' thinking in arithmetic. In Arithmetic 1949 (compiled & edited by G. T. Buswell & Maurice L. Hartung). Supplementary Educational Monographs, No. 70. Chicago: University of Chicago Press, November 1949. (Pp. 55-63)

19. Dawson, Dan T., & Ruddell, Arden K. An experimental approach to the division idea. Arithmetic Teacher 2: 6-9; February 1955.

20. Gunderson, Agnes G. Thought patterns of children in learning multiplication and division. Elementary School Journal 55: 453-461; April 1955.

21. Weaver, J. Fred. Big dividends from little interviews. Arithmetic Teacher 2: 40-47; April 1955.

22. Buswell, Guy T., & Kersh, Bert Y. Patterns of thinking in solving problems. University of California Publications in Education, Vol. 12, No. 2. Berkeley: University of California Press, 1956. (Pp. 63-148)

23. Gibb, E. Glenadine. Children's thinking in the process of subtraction. Journal of Experimental Education 25: 71-80; September 1956.

24. Van Engen, Henry, & Gibb, E. Glenadine. General mental functions associated with division. Educational Service Studies, No. 2. Cedar Falls: Iowa State Teachers College, 1956.

25. Olander, Herbert T., & Brown, Betty Irene. A research in mental arithmetic involving subtraction. Journal of Educational Research 53: 97-102; November 1959.

(Continued)

26. Ruddell, Arden K. Levels of difficulty in division. Arithmetic Teacher 6: 97-99; March 1959.

27. Pace, Angela. Understanding and the ability to solve problems. Arithmetic Teacher 8: 226-233; May 1961.

28. Brownell, William A. Arithmetical abstractions: Progress toward maturity of concepts under differing programs of instruction. Arithmetic Teacher 11: 547-556; December 1964.

30. Brace, Alex, & Nelson, L. Doyal. The preschool child's concept of number. Arithmetic Teacher 12: 126-133; February 1965.

31. Gray, Roland F. An experiment in the teaching of introductory multiplication. Arithmetic Teacher 12: 199-203; March 1965.

32. Gray, Roland F. An approach to evaluating arithmetic understandings. Arithmetic Teacher 13: 187-191; March 1966.

33. Hervey, Margaret A. Children's responses to two types of multiplication problems. Arithmetic Teacher 13: 288-292; April 1966.

34. Brownell, William A. Conceptual maturity in arithmetic under differing systems of instruction. Elementary School Journal 69: 151-163; December 1968.

35. Kilpatrick, Jeremy. Analyzing the solution of word problems in mathematics: An exploratory study. Unpublished doctoral dissertation, Stanford University, 1967. (Dissertation Abstracts 28: 4380A; May 1968.

36. Rea, Robert E., & Reys, Robert A. Mathematical competencies of entering kindergarteners. Arithmetic Teacher 17: 65-74; January 1970.

37. Reys, Robert E., & Rea, Robert E. The Comprehensive Mathematics Inventory: An experimental instrument for assessing the mathematical competencies of children entering school. Journal for Research in Mathematics Education 1: 180-186; May 1970.

38. Lankford, Francis G., Jr. Some computational strategies of seventh grade pupils. Final Report, Project No. 2-C-013. U.S. Department of Health, Education, & Welfare, Office of Education, National Center for Educational Research & Development, and Center for Advanced Study, University of Virginia. October 1972.

39. Grouws, Douglas A. Open sentences: Some instructional considerations from research. Arithmetic Teacher 19: 595-599; November 1972.

(Continued)

40. Pereira, W. A study relating girls' overt, verbalization performance, retention, rules and strategies as they learn a mathematical structure based on elements of a potential theory which relates thinking language and learning. (Technical Report No. 293.) 1973, Wisconsin Research and Development Center.

41. Bechtel, R. D. & Weaver, J. F. A preliminary investigation of the ability of second-grade children to solve division problems presented in a physical-object mode and in different partitive-measurement sequences. (Working Paper No. 194.) 1976, Wisconsin Research and Development Center.

J.F. Weaver
University of Wisconsin-Madison
18 April 1973
Revised April 1976

A Clinical Model for Diagnosing Mathematical Deficiencies, $(MD)^2$
Incorporating Educational Cognitive Style

William R. Speer
Bowling Green State University

Introduction

One of the most significant movements toward improved instruction that has been made in recent years is that of incorporating into the instructional setting plans for diagnosing learning difficulties and detecting needs for preventive and corrective teaching (Engelhardt, 1974, 1976; Irons, 1976). In fact, the term "diagnosis" and its variety of meanings have become part of the vocabulary of mathematics educators, with some notable exceptions, at all levels of instruction (Crenson, 1976). As an example, consider the Guidelines for the Preparation of Teachers of Mathematics (1973) published by the Commission on Preservice Education of Teachers of Mathematics of the National Council of Teachers of Mathematics which put forth recommendations that included the attainment of competencies in "recognizing stages of cognitive, affective, and psychomotor development in children and individual differences between children as these differences pertain to the learning of mathematics," (p. 15) as well as in diagnosing and prescribing "remedies for common disabilities in the learning of mathematics and to know what tools and techniques are available to help with diagnosis and correction" (p. 15).

Nevertheless, when the literature dealing with the diagnosis of mathematical difficulties is examined, one finds "a paucity of relevant research and a lack of substantive contributions" (Cawley, 1975, p. 12). "Research, in the main, has been perfunctory at best and dissemination almost non-existent" (Crenson, 1976, p. 1). Suiler (1970), in the National Council of Teachers of Mathematics Thirty-Second Yearbook, reported a study which categorized research studies in mathematics education made in the United States between 1880 and 1963. The category focusing on diagnosis contained a total of only 103 entries, which in graphical form, reflects a bimodal distribution with the modes appearing in the late 1920's and again in the early 1940's. A similar study by Suydam (1970) located a total of 107 studies focusing on diagnosis and/or remediation for the period ranging from 1900 to 1965. Wilson, reporting on the nature of research in mathematics education, stated that any search of available resources

> readily reveals that the type of work I've called clinical intervention [guiding efforts in diagnosis

> and treatment of children's learning difficulties in mathematics/ is either not recognized as research, or is depreciated as research, and in some circles even deprecated as research. (Wilson, 1973, pp. 1-2)

Several recent calls for needed research appear to be having an impact with respect to generating new interest in research and application of diagnostic models (Gray, 1972; National Advisory Committee on Mathematical Education, 1975; National Assessment of Educational Progress, 1973; National Conference on Needed Research in Mathematics Education, 1967; National Conference on Remedial Mathematics, 1974, 1975, 1976). A cursory examination of recent National Council of Teachers of Mathematics regional and national meetings indicates expanding attention to the diagnostic-remedial arena. State governments also provide impetus for further gains by the passage of laws such as Section 230.2311 of the Florida Statutes (1975) which states, in part, that each school district's program shall include an individualized diagnostic approach to instruction.

> Probably the '70s will witness a kindling of efforts by a national cross-sectional group of individuals from diverse institutions representing isolated efforts in diagnosis in clinics and teacher training programs. The most identifiable catalyst in this movement is the effort of Jim Heddens at Kent State University who, through the help of the KEDS General Assistance Center, has organized three national conferences with diagnostic themes. (Underhill, 1976, p. 1)

Callahan in his summary paper of the first National Conference on Remedial Mathematics, stated:

> Many participants commented on the level of knowledge that exists in regard to diagnostic-remedial procedures in mathematics. It would seem fair to say that there is not a great deal of systematic, accumulated knowledge. Some smatherings of research evidence and some sensitive and insightful thoughts on the subject exist. Some isolated individuals at various points in time have attempted to pull together some of the research and thoughts, but the level of scientific knowledge regarding the diagnostic-remedial episode in mathematics is not great. (Callahan, 1974, p. 3)

Many examples have been used to cite this apparent lack of concern for providing a sound knowledge base for research on diagnosis (National Conference on Needed Research in Mathematics Education, 1967). What is perhaps more important is why this state of affairs exists. Engelhardt (1974) proposed three major factors as possible explanations. These are (1) "an insufficient understanding of the learning process," (2) "the disjointed nature of most research efforts," and (3) "the apparent lack of theoretical models for mathematical remediation, models which would identify sets of variables to provide a common focus for research" (pp. 2-3). These three factors provide the framework for this paper.

Purpose

The purpose of this paper is to describe a theoretical model for the clinical diagnosis of mathematical deficiencies--a model which is compatible with and built upon research and supported by theory. The research base for a study of this nature is not confirmative (i.e., one which is experimental) but rather it is generative (i.e., one which generates hypotheses) and analytic-synthetic (i.e., one which constructs guiding models and explanatory theories) (Wilson, 1973).

Specifically, the model focuses on three considerations, namely, the identification of specific mathematical deficiencies, the identification of mathematical cognitive style, and the identification of general educational cognitive style. The model is designed and developed primarily for use in a clinical setting.

Identification of specific mathematical deficiencies refers to the process of determining, by various forms of examination the specific nature and circumstances of a more general, suspected deficiency. For the purposes of this paper the diagnosis of mathematical deficiencies is limited to the concept clusters represented by the Kent State University Mathematics Checklist (1975). It is important to note, however, that the possibility exists for the extension of this model into other areas and levels of mathematics. Indeed, if this model is to provide any significant contribution to the field of diagnosis it .ust have this property. "The creative aspects of mathematics, skill in searching out mathematical patterns, non-routine problem solving, and the study of functional relationships must be considered as carefully as the skill aspects" (Riedesel, 1974, p. 1).

Mathematical cognitive style is viewed as the manner in which an individual operates on the concrete--representative--abstract hierarchy in light of mediating variables (i.e., visual, auditory, tactile) encountered in both the mode of presentation and type of desired student response.

Educational cognitive style, as it is used in this paper has its foundation in the work done by Hill (1968, 1972, 1974).

Informally, educational cognitive style is the way in which an individual takes on meaning, the way in which an individual perceives his surroundings, the way in which an individual can master an educational task most readily (Hill, 1974).

The purpose of the paper, then, is to combine the three factors of mathematical diagnosis, mathematical cognitive style, and educational cognitive style into a model and instrument which can be used to generate a map of the individual student. This map can then be used in a prescriptive sense to determine the most probable means of providing successful corrective teaching.

Theoretical Framework

One of the assumptions of this paper is that "man is not content with biological satisfactions alone, but rather he continually seeks meaning" (Hill, 1974, p. 2). The implication is, of course, that we are inherently curious and constantly searching for reasons and explanations which give meaning to our environment as we perceive it. The fruits of our search for knowledge and understanding, in this instance, are represented by a profusion of theories and models which attempt to explain various aspects of the educational arena.

Some theories are formulated on the basis of painstaking observation and categorization--what has been called "confirmative research" (Wilson, 1973). The foundation of such theories lies in the gathering of experimental data which is then used to develop classification schemata. In mathematics, for example, this is akin to the inductive reasoning process used by elementary school children who generalize, after several experiments, that the sum of two even numbers is an even number.

Other theories are predicated on the basis of rational analysis. That is to say, a general theory is put forth, based on certain assumptions, and then "experiments" or observations are made to see whether or not they fit the theory. Essentially, this is the process called "empirical mapping" by Hill (1974), or "generative research" by Wilson (1973). An example of such a process can be found in the works of Piaget (1952).

An important characteristic of confirmative or experimental research has been described by Mouly.

> Experimentation, whether in education or any other field, rests on the assumption that there exist invariant relationships between certain antecedents and certain consequents so that, provided a given set of conditions prevail, if one does this, that will follow. (Mouly, 1968, p. 6).

But, exceptions to the "rule" do exist. In education it is particularly easy to rationalize these exceptions into oblivion by calling on some unusual circumstance as the cause of the unexpected exception. In many instances, of course, there may indeed be some "unusual circumstance" acting. The multitude of variables involved in many educational studies surely increase our chances of overlooking an important factor.

> The difference between what was planned and actually occurred is considerable. Furthermore, this difference is intended. Corrective instructional events are not mechanistic routines to be blindly followed. Real events grow, change and develop as the human beings involved in the event interact. (Romberg, 1976, p. 3)

Nevertheless, this technique of rationalizing identified contradictions sometimes amounts to little more than burying our heads in the sand. About the best that can be said is that experimentation in education has produced a number of generalizations, principles, and laws which are valid under certain stated conditions (Van Dalen, 1966; Romberg, 1976).

Information gathered in the name of rational analysis, personnal experience, intuition, or opinion has the same characteristic. A great deal of the "knowledge" we possess relative to our students has been derived through informal observation, and consequently generalizations, principles, and laws formulated in this manner are valid under certain stated conditions. It appears, at least to the author, that when we deal with the "less tangible," less experimentally oriented aspects of education there is little difference between the experimental and experiential approaches, except perhaps that in the experiential model we do not try to rationalize inconsistencies into oblivion (Nunnally, 1975).

> Instead of making generalizations the ruling consideration in our research, I suggest that we reverse our priorities. An observer collecting data in one particular situation is in a position to appraise a practice or proposition in that setting, observing effects in context. In trying to describe and account for what happened, he will give attention to whatever variables were controlled, but he will give equally careful attention to uncontrolled conditions, to personal characteristics, and to events that occurred during treatment and measurement. As he goes from situation to situation his first task is to describe and

> interpret the effect anew in each locale, perhaps taking into account factors that were unique to that locale or series of events. (Cronbach, 1975, p. 117)

This is essentially what Romberg (1976) called "process evaluation"; "process" in the sense that the focus is directed toward actions as opposed to outcomes, and "evaluation" in the sense that it is a question raising search rather than conclusion drawing research.

Wilson summarized the position of research of this nature in the following manner.

> The reasons for neglecting the systematic development and use of clinical intervention as a type of research are varied and have deep historical roots. Of those which are most often mentioned to me, the most common are that "such studies aren't 'rigorous'", "they don't use controls", "they use procedures and data that don't lend themselves to the analyses of inferential statistics", "you can't generalize from one child or even a small group so studied", "there is no way to replicate", etc.
>
> Such criticisms are based on criteria appropriate to experimental research. For studies that are labeled as experiments _and_ intended by the researcher to fulfill the purposes of experimental research, such criticisms are, of course, accurate and fully justified.
>
> That studies which do not claim to be experiments are also criticized on these grounds--if only implicity--attest to the eminent position the criteria of excellence in experimental research have attained in our community. The high esteem we have for correctly designed and executed experimental research is fully justified--for the purposes to which experimental research is suited. A somewhat comparable and justified esteem is held for sound correlational studies. But is it possible this esteem has obscured our clear recognition of the potential value of other kinds of research? In turn has this inhibited our efforts to improve other kinds of research? (Wilson, 1973, p. 2)

One often hears, particularly in educational contexts, that theory and practice are not even related, let alone isomorphic

(Newsome, 1964). Others, including the author, do not agree with this position. Reys and Post (1973) stated that, "Facts play a central role in the development of theory, and the theory subsequently provides a systematic interpretation of the general area to which the facts are related" (p. 16). Hill stated that

> It is a serious mistake to think of a realm of theory that is separate and different from the realm of fact. It would be reasonable to say either that facts represent one kind of theory or that theories represent one kind of fact, but most reasonable to say that fact and theory represent different degrees of what is basically a single process. (Hill, 1963, p. 23)

Burns (1962) suggested that the nature of the relation between theory and practice can best be described by "pragmatic implication" which he defines as "a rational person with certain beliefs relevant to certain kinds of situations generally acts in accord with those beliefs" (p. 54). Guttchen (1966) took exception to Burns' approach, mainly because the terms "rational" and "relevant" seem to leave too much room for different interpretation. He did, however, allow for the possibilities of action according to pragmatic implication in such areas as medicine or engineering.

Gowin (1963) posed the view that theory is a blending of logic and facts and that it is best seen as a guide to thought and inquiry. In other words, Gowin points to the true nature of the relationship between the strategies of studies done in an experimental vein and those founded on rational analysis. Perkinson (1964) preferred to couch educational theory in terms of a strategy--what Gowin (1964) referred to as a flow chart to guide experimentation.

Clements (1962) described two basic types of theories, namely "prescriptive" and "descriptive." Of these two, the former represents what is generally called "educational theory." According to Soltis (1968), "Descriptive theories are adequate when they allow for accurate predictions and little, if any, educational theory is now of this sort" (p. 85).

Two opposing views are presented by Reys and Post (1973) and Newsome (1964). Reys and Post contended that, "Ideally, theories should provide insight to both theorist and practitioner concerned with a common area of investigation" (p. 17). Newsome, on the other hand, argued that the relationship between theory and practice is negligible. Theory does allow for better understanding of practical situations, according to Newsome, but it does not describe a set of logical processes to be applied to any given situation.

Whatever the organizing foundation for theory may be, the development of the general principles and laws contained within the structure of the theory provide us with the means to at least attempt to predict and control events in our surroundings. The degree to which we have predictability and control is dependent upon several factors including "goodness of fit (how much agreement there is between the model and the phenomena it is attempting to describe), "relevance" (the degree to which the theory matches other characterizations, particularly those that have "checked out") and "fruitfulness" (the development of checkable characterizations beyond those already in existence).

Golladay, DeVault, Fox, and Skuldt (1975) identified problems in empirical research in mathematics education. It is argued that it is often difficult to choose "appropriate conceptualizations and measures for a variety of phenomena" (p. 159) found in the study of individuals and individual differences. Further, "models of more traditional, structured educational experiences are not appropriate for examining the greater variety of opportunities and experiences characteristic of most individualized programs" (p. 160). A call is made for the use of paradigms, i.e., descriptors, to identify categories and relations with the intent of organizing observed data. They pointed to the successful use, by the scientific community, of paradigms but quickly draw from a study by Apple (1973) which indicated that while educators often employ paradigms they are seldom specifically stated.

Golladay et al. also pointed to the problem of reliability in studies which focus on the individual. The major cause is the complexity of events which are presented to the observer. They concluded that

> It may well be inappropriate to search for traditional methods for testing the reliability of information when the program being studied departs from traditional patterns ... and information is gathered in a manner different from that of traditional research designs. (Golladay et al., 1975, p. 168)

Since the present paper is concerned with the development of a model for use in diagnosing an individual's mathematical deficiencies and identifying individual educational cognitive style it seems that the notions presented in the previous pages, are relevant. However, an additional word of caution is necessary. An inherent danger of developing an illustrative model (or theory, for that matter) is oversimplification to the point that distortion makes the model (or theory) useless. On the other hand, presenting a model (or theory) which incorporates all of the complexities of the situation under study runs the danger of being too copious to allow for practical application.

Nature of Diagnostic Models

Diagnostic models are of three varieties in terms of the setting in which the diagnosis is to take place, namely, models that are classroom oriented, models designed for clinical use, and models which may be applied to either environment. The primary emphasis in this paper is on clinical models since the $(MD)^2$ model is of this type.

Diagnostic models can also be differentiated on the basis of their assumptions concerning the purpose of diagnosis and its corresponding methodologies.

One type of model, the ability training model, has as its purpose the identification of learner capabilities which, when identified, may be used to prescribe corrective teaching (Uprichard, Baker, Dinkel & Archer, 1975). Thus, ability training may be roughly equated with aptitude-treatment-interaction (ATI) which "seeks to provide a basis for employing differential treatments in order to exploit the cognitive preferences displayed by different individuals for differing content or mode of instruction" (Hancock, 1975, p. 37).

Some objections to the use of the ability training model that have been cited include: (1) the nature of the operational definitions used, (2) the difficulty of incorporating ATI findings in the instructional setting, (3) the instability of reliability and validity measures of instruments used to gather data, and (4) the lack of research which supports the notion that remediating weaknesses in cognitive preferences leads to increased performance in the classroom (Stiglmeier, 1972; Uprichard et al., 1975).

Each of these objections can be countered by referring to the available literature. For example, objections one and two can be dispelled by providing a scientific framework for education and the accompanying means of implementing this framework in educational settings (Hill, 1974). Objection three has been discussed at great length (Hill, 1973) and poses no problem provided results are properly interpreted. Concerning objection four, it has been stated that, "the lack of productivity in this area has been ascribed to inadequacies in research design and general methodology" (Cunningham, 1975, p. 171). However, there now exists an abundance of research which refutes the objection that application of ability training techniques does not increase achievement. Not only has it been shown that remediating weakness in the child's cognitive preferences leads to increased performance in learning situations but also it has been shown that remediation based on utilizing the child's cognitive preference strengths has similar effects (Radike. 1973).

The second theoretical model for diagnosis is known as the task-analysis model. This content oriented method consists of "analyzing a learning task into a hierarchy of subordinate tasks, diagnosing the pupils' mastery of the subordinate tasks [illegible] giving instruction in the specific subordinate tasks not master[illegible]y the learner" (Callahan & Robinson, 1973, p. 579). According to Uprichard et al. (1975), in task analysis "the emphasis is on component skills and their integration into complex terminal tasks rather than the processes that presumably underlie the development of specific tasks" (p. 2).

Identified criticisms of this model include: (1) the content-orientation may cause the diagnostician to overlook important factors in the student, (2) the task analysis of certain subjects is difficult, (3) the validation of hierarchies is a difficult process, and (4) the prescriptive philosophy of task analysis tends to be founded on changes in the curriculum, for the most part ignoring changes which reflect analysis of the learners' cognitive style (Uprichard et al., 1975).

Despite these criticisms several studies have pointed to the value of the task analysis model. For example, Uprichard et al. (1975) stated that "the task analysis model has appeal for /mathematics/ educators since the structure of the discipline aids in the building of hierarchical relationships.... /it/ is valuable in that the diagnostic findings rely on fewer undefined and unvalidated assumptions" (p. 2). Additional benefits of the task analysis model have been suggested.

> One conjecture is that a procedure of diagnosis and instruction based on a hierarchical analysis or subordinate tasks is an effective procedure for students' learning of a mathematical task.... Another conjecture is that where the task analysis procedure is used in the teaching of a mathematical task the incidence of underachievement ... will significantly decrease.... In summary, the task-analysis procedure when combined with meaningful mastery learning of the subordinate tasks in a hierarchy seems quite effective in learning a mathematical task. (Callahan & Robinson, 1973, pp. 583-584)

It should be noted at this point that the model described in this paper is a synthesis of both the ability training and the task analysis theories of diagnosis. The Model for Diagnosing Mathematical Deficiencies is designed to provide information about student style as well as content deficiencies. The corrective teaching procedures suggested by the model include not only revisions in content but also in modes of instruction to better match the individual's unique cognitive preferences.

Another method of differentiating between diagnostic models is by identifying their research base, specifically, by referring to the manner in which the various theories of diagnosis are developed.

Wilson (1973) described three major classes of research. The first, confirmative research, is experimental in nature and centers on "activities designed to assess the truth of probable hypotheses" (p. 11). The second, analytic-synthetic research, deals with "activities involved in the development of guiding paradigms and explanatory theories" (p. 11). Third, generative research, is based on "activities consistent with the postulates of science designed to generate hypotheses with an a priori probability" (p. 10). Generative research is further described by Wilson in terms of two subtheories: (1) normative research with "activities designed to generate hypotheses concerning facts and those connections between facts which exist in nature" (p. 16). and (2) clinical intervention research which involves "activities designed to generate hypotheses on those connections between facts which might be brought into nature by some intervention" (p. 16).

The model described in the present paper, the clinical Model for Diagnosing Mathematical Deficiencies--(MD)^2, has a research base which is both generative and analytic-synthetic. Aptitude-treatment-interaction theories are also founded on the generative research approach by nature of their study of the relevant processes engaged in by students in learning situations with the intent of generating hypotheses concerning these processes (Wittrock, 1974). Much of Piaget's work has centered on this same form of research. Piagetian-type research has impelled us to

> take a fresh look at our field and to ask a host of new questions concerning the nature of developmental stages and of developmental processes generally, as well as of the kind of research approaches which the study of these problems demands. In so doing it has helped us appreciate the important place of systematic theory in an area of developmental research, essentially comparative in nature, which has not always been noted for its theoretical sophistication. On the other hand, the theoretical significance of research inspired by Piaget's ideas does not prevent it from having direct and important relevance

> for the resolution of practical questions of pedagogy and educational practice. (Wohlwill, 1968, p. 446)

To summarize, diagnostic models can be differentiated in terms of whether they focus on classroom and/or clinical procedures, on whether they are ability-training or task analysis oriented, and on whether they have a research base which is confirmative, generative, or analytic-synthetic. These different approaches to the diagnostic-prescriptive arena do not necessarily reflect differences on what diagnosis is, but rather on how diagnosis is to be carried out.

Nature of Educational Cognitive Style

Before discussing the $(MD)^2$ model, it is necessary to examine one of its components in some detail. This component, educational cognitive style, is not content oriented toward mathematics but its usefulness in the prescriptive stage will readily become evident.

> There will be a great change made in the first and foremost and continuing business of society: the education and training of the young. The development of the mind of the child will come to rest in the knowledge and skills of the biochemist, the pharmacologist, and neurologist, and psychologist, and educator. And there will be a new expert abroad in the land--the psychoneurobiochemeducator. (Krech, 1969, p. 374)

While the "new expert" that Krech refers to may still be somewhat futuristic, advances have been made to develop a more scientific, but not less humanistic, framework for education. The most notable of these efforts falls under the auspices of the American Educational Sciences Association and its founder, Joseph Hill. Since its inception in 1971 (Hill's first published work in this area was in 1966) its membership has grown to over 250 (AESA Membership Directory, 1975) and a recent AESA bibliography (Berry, Sutton, & McBeth, 1975) included over 300 entries on various aspects of the Educational Sciences. Also, the educational science of cognitive style has had considerable impact on educational programs at all levels and has been adopted by many school systems. It is recognized that shear numbers are no indication of the value of any organization or cause. This data is provided only for the reason that it dispels any thought that knowledge of, and development in, the Educational Sciences is limited to a select few.

The following quotes suggest the rationale for the development of the Educational Sciences.

> American education presents to the public view the spectacle of a house divided against itself. One needs only to persue the back copies of educational journals to see how the battle has raged, and to observe the disarray of the schools. Each conflicting point of view finds its advocates.
>
> It is obvious that the confusion and disarray in education arises from the lack of commonly agreed upon goals, practices, and definitions. In other words, instead of having a common framework and a common language, educators have developed an amorphous collection of ideas, concepts and methods from a variety of other disciplines. (Radike, 1973, Introduction)

> Without a framework of 'language', the vast field of human activity called 'education' does not readily lend itself to meaningful description or definition. At the present time, the universe of discourse associated with education lacks precision beyond that found at the levels of common sense and daily journalism. The difficulty with such language is not that it fails to provide a form of communication, but that the possibilities of misunderstanding are great and the probability of relatively precise discriminations and predictions is small. (Hill, 1968, p. 1)

> Many educational terms do not have clearly assigned and commonly understood meanings, when words such as 'democracy', 'education', 'curriculum', and 'discipline' are used by different workers in the field, they may stand for slightly or radically different things. In contrast, the technical terms in the exact sciences such as meter, ampere, lightyear, and calorie are instruments of great exactitude. (Van Dalen, 1966, p. 200)

These exact sciences referred to by Van Dalen can be equated to what Hill refers to as "fundamental disciplines" (Hill, 1974).

> Fundamental disciplines are bodies of knowledge generated by communities of scholars that produce pure and distinctive forms of information about phenomena which they study. Biology, history, art, psychology, and mathematics are examples of fundamental disciplines.

> Complementing the fundamental disciplines are the applied or derivative fields of knowledge. These bodies of information are generated by practioners who deal with practical considerations of the human condition. Medicine, pharmacy, engineering, and law are examples of applied fields of knowledge. (Hill, 1974, p. 1)

Education is not a fundamental discipline but instead is an applied or derivative field. The Educational Sciences represent an attempt to describe a conceptual framework for education that is as precise and definite as that found in other applied fields. According to Hill,

> With the development of the Educational Sciences, the solutions of problems and explanations of phenomena are facilitated, and educational problems accuring to inadequate communication, misinterpretation of information, and fragmentation of effort are alleviated. (Hill, 1974, p. 1)

Presently there are seven educational sciences. These include: (1) symbols and their meanings, (2) cultural determinants of the meanings of symbols, (3) modalities of inference, (4) biochemical and electrophysiological aspects of memory, (5) cognitive style of individuals, (6) teaching styles, administrative styles and counseling styles, and (7) systemic analysis decision-making. The fifth educational science, cognitive style of individuals, includes the first three educational sciences (the fourth is not sufficiently developed at this point). Therefore, for the purposes of this paper the discussion will center on the educational sciences of symbols and their meanings, cultural determinants, and modalities of inference, i.e., educational cognitive style.

Classroom teachers have long been aware that students come to know what they know in their own unique way. Until recently, however, there was no established framework for teachers to analyze the learning habits of their students and match them with the "most appropriate" mode of instruction. Educational cognitive style provides such a framework.

Briefly, educational cognitive style is a means of identifying the ways in which an individual perceives and reacts to the environment. An individuals' cognitive style is the way a student tends to seek meaning and the manner in which information is personally filtered. Cognitive styles are influenced by the ways in which individuals derive meaning from symbols related to their personal experiences and the world about them; the influences of family, friends, and their own individuality on these meanings; and the kind of reasoning processes used to derive these meanings.

A BRIEF GUIDE TO COGNITIVE STYLE MAPPING
Symbols and Their Meanings

Reprinted with permission from Joseph Hill (1976)

Two types of symbols, theoretical (e.g., words and numbers) and qualitative (e.g., code data) are basic to the acquisition of knowledge and meaning. Theoretical symbols differ from qualitative symbols in that the theoretical symbols present to the awareness of the individual something different from that which the symbols are. Words and numbers are examples of theoretical symbols. Qualitative symbols are those symbols which present and then represent to the awarness of the individual that which the symbol is. (Feelings, commitments and values are some examples of the meanings conveyed by the qualitative symbols.)

There are four Theoretical Symbols:

T(VL) Theoretical Visual Linguistic — ability to find meaning from words you see. A major in this area indicates someone who reads with a better than average degree of comprehension.

T(AL) Theoretical Auditory Linguistic — ability to acquire meaning through hearing spoken words.

T(VQ) Theoretical Visual Quantitative — ability to acquire meaning in terms of numerical symbols, relationships, and measurements.

T(AQ) Theoretical Auditory Quantitative — ability to find meaning in terms of numerical symbols, relationships and measurements that are spoken.

Meanings for qualitative symbols are derived from three sources: 1) sensory stimuli; 2) cultural codes (games); and 3) programmatic effects of objects which convey an almost automatic impression of a definite series of images, scenes, events or operations. At the present time, there are 20 qualitative symbols included in the "symbolic" set; five of them associated with sensory stimuli, five that are programmatic in nature, and ten associated with cultural codes.

The five qualitative symbols associated with sensory stimuli are:

Q(A) Qualitative Auditory — ability to perceive meaning through the sense of hearing. A major in this area indicates ability to distinguish between sounds, tones of music, and other purely sonic sensations.

Q(O) Qualitative Olfactory — ability to perceive meaning through the sense of smell.

Q(S) Qualitative Savory — ability to perceive meaning by the sense of taste. Chefs should have highly developed qualitative olfactory and savory abilities.

Q(T) Qualitative Tactile — ability to perceive meaning by the sense of touch, temperature, and pain.

Q(V) Qualitative Visual — ability to perceive meaning through sight.

The qualitative symbols that are programmatic in nature are:

Q(P) Qualitative Proprioceptive — ability to synthesize a number of symbolic mediations into a performance demanding monitoring of a complex task (e.g., playing a musical instrument, typewriting); or into an immediate awareness of a possible set of interrelationships between symbolic mediations, i.e., dealing with "signs."

Q(PD) Qualitative Proprioceptive Dextral — a predominance of right-eyed, right-handed and right-footed tendencies (a typically right-handed person) while synthesizing a number of symbolic mediations into a performance demanding monitoring of a complex task (e.g., playing a musical intrument, typewriting).

Q(PK) Qualitative Proprioceptive Kinematics — ability to synthesize a number of symbolic mediations into a performance demanding the monitoring of a complex physical activity involving motion.

Q(PS) Qualitative Proprioceptive Sinistral — a predominance of left-eyed, left-handed and left-footed tendencies (a typically left-handed person) while synthesizing a number of symbolic mediations into a performance demanding monitoring of a complex task (e.g., playing a musical instrument, typewriting).

Q–PTM) Qualitative Proprioceptive Temporal — ability to synthesize a number of symbolic mediations into a performance demanding the monitoring of a complex physical activity involving timing.

The remaining ten qualitative symbols associated with cultural codes are defined as:

Q(CEM) Qualitative Code Empathetic — sensitivity to the feelings of others; ability to put yourself in another person's place and see things from his point of view.

Q(CES) Qualitative Code Esthetic — ability to enjoy the beauty of an object or an idea. Beauty in surroundings or a well-turned phrase are appreciated by a person possessing a major strength in this area.

Q(CET) Qualitative Code Ethic — commitment to a set of values, a group of principles, obligations and/or duties.

Q(CH) Qualitative Code Histrionic — ability to exhibit a deliberate behavior, or play a role to produce some particular effect on other persons. This type of person knows how to fulfill role expectations.

Q(CK) Qualitative Code Kinesics — ability to understand, and to communicate by, non-linguistic functions such as facial expressions and motions of the body (e.g., smiles and gestures).

Q(CKH) Qualitative Code Kinesthetic — ability to perform motor skills, or effect muscular coordination according to a recommended, or acceptable, form (e.g., bowling according to form, or golfing).

Q(CP) Qualitative Code Proxemics — ability to judge the physical and social distance that the other person would permit, between oneself and that other person.

Q(CS) Qualitative Code Synnoetics — personal knowledge of oneself.

Q(CT) Qualitative Code Transactional — ability to maintain a positive communicative interaction which significantly influences the goals of the persons involved in that interaction (e.g., salesmanship).

Q(CTM) Qualitative Code Temporal — ability to respond or behave according to time expectations imposed on an activity by members in the role-set associated with that activity.

Cultural Determinants

There are three cultural determinants of the meaning of symbols: 1) individuality, 2) associates; and 3) family. It is through these "determinants" that cultural influences are brought to bear by the individual on the meanings of symbols

F — Family   I — Individual   A — Associates

Modalities of Inference

The forms of inference the individual uses in the process of deriving meaning:

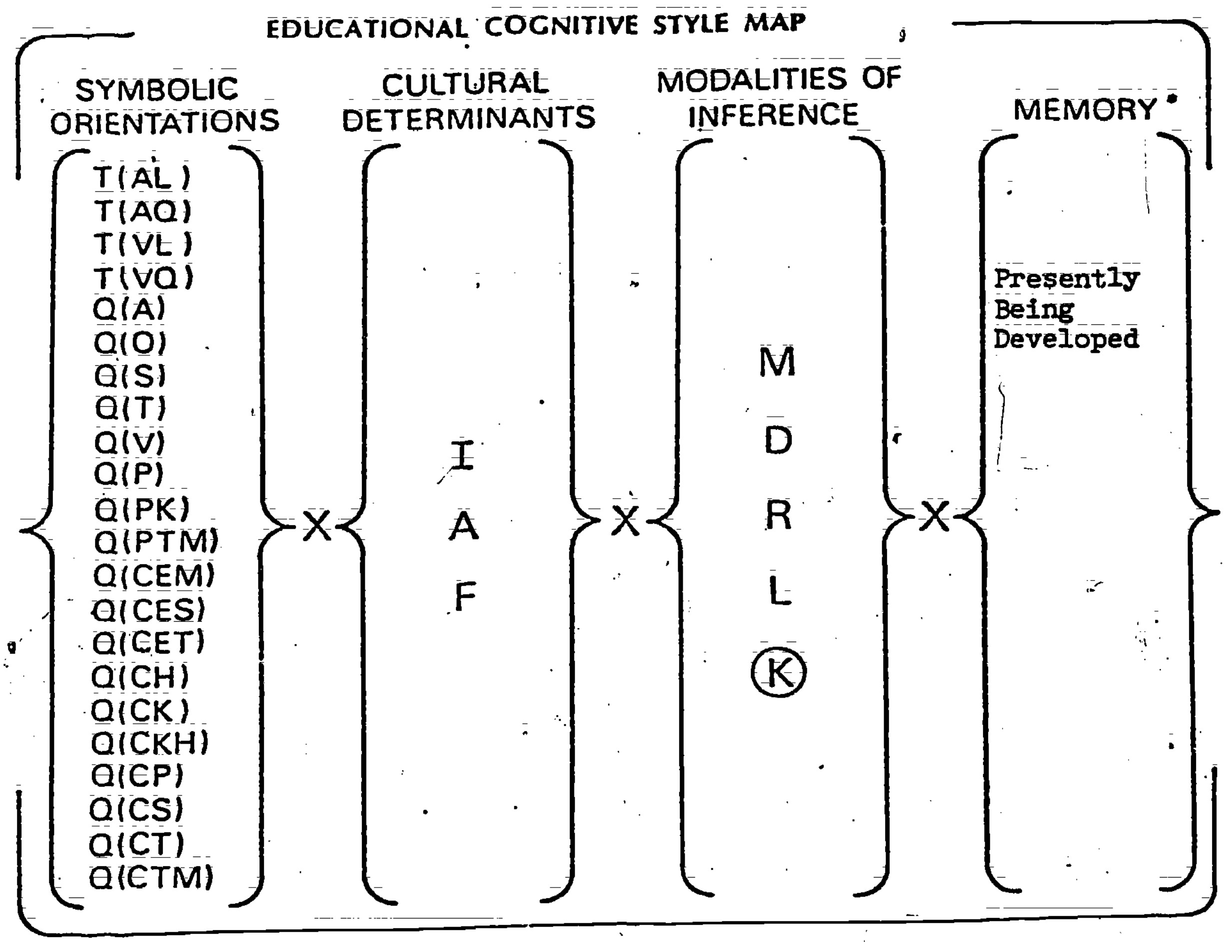
M Magnitude — a form of "categorical reasoning" that utilizes norms or categorical classifications as the basis for accepting or rejecting an advanced hypothesis. Persons who need to define things in order to understand them reflect this modality.

D Difference — This pattern suggests a tendency to reason in terms of one-to-one contrasts or comparisons of selected characteristics or measurements. Artists often possess this modality as do creative writers and musicians.

R Relationship — this modality indicates the ability to synthesize a number of dimensions or incidents into a unified meaning, or through analysis of a situation to discover its component parts. Psychiatrists frequently employ the modality of relationship in the process of psychoanalyzing a client.

L Appraisal — is the modality of inference employed by an individual who uses all three of the modalities noted above (M, D, and R), giving equal weight to each in his reasoning process. Individuals who employ this modality tend to analyze, question, or, in effect, appraise that which is under consideration in the process of drawing a probability conclusion.

K Deductive — indicates deductive reasoning, or the form of logical proof used in geometry or that employed in syllogistic reasoning.



By utilizing the techniques of observation, interview, and preference testing, a diagnostician can gather and assemble data on the elements of each of the three major components of educational cognitive style to form a profile, or "cognitive style map" of the individual student. These elements may appear as major orientations (if the element score occurs in the 50th-90th percentile of a distribution of scores for that element), minor orientations, denoted by a prime (if the element score occurs in the 26th-49th percentile), or negligible orientations (if the element score is at or below the 25th percentile).

The testing procedure used to arrive at a cognitive style map has received considerable discussion elsewhere (Radike, 1973), and consequently a further detailed treatment seems inappropriate. Suffice it to say that the diagnostician may use any one, or all, of the three methods: (1) observation, (2) interview, and (3) preference testing.

> The mapping of cognitive styles is mainly <u>empirical</u> in nature, and as such, is dependent upon the judgments of persons (diagnosticians) ... The cognitive style of an individual cannot be empirically mapped without considering: (1) the level of educational development of the person, (2) the general symbolic conditions of educational tasks he will be called upon to accomplish, (3) certain antecedents (e.g., family) to his present state of development, and (4) the appropriateness of the elements under consideration for the conditions under which the educational tasks must be completed. (H'll, 1970, p. 7)

For those readers that are interested, Radike (1973) presents a valuable summary of the process of empirical mapping.

The educational cognitive style model is similar in some respects to the Task-Process Integration Model (Uprichard et al., 1975) however, it is content-free and considerably more global in its approach to student's learning style. Educational cognitive style diagnosing can best be described as a combination of classroom and clinical procedures and is clearly an ability training model although once the learner diagnosis is complete a form of task-analysis is used in determining the symbolic orientation of instructional resources.

The Model for Diagnosing Mathematical Deficiencies--$(MD)^2$

There has been considerable research focusing on the traits of successful mathematics students (Shuart, 1970). The results, while tending to be inconclusive, do suggest that several factors need to be given greater attention than they may have been given in the past. A partial list of "identified" traits, summarized from Shuert (1970), is found below. [Note that the traits are grouped, roughly, as they relate to symbolic orientations, cultural determinants, and modalities of inference as found in educational cognitive style (Hill, 1974).]

The successful mathematics student

has high general and reflective intelligence
prefers objective, non-personal symbolism
is high in verbal ability and comprehension
is highly competitive
possesses authoritarian attitudes
tends to be insecure and sensitive
tends to avoid social and interpersonal issues
rates high on self-acceptance
is anxious
is concerned with "abstract" beliefs

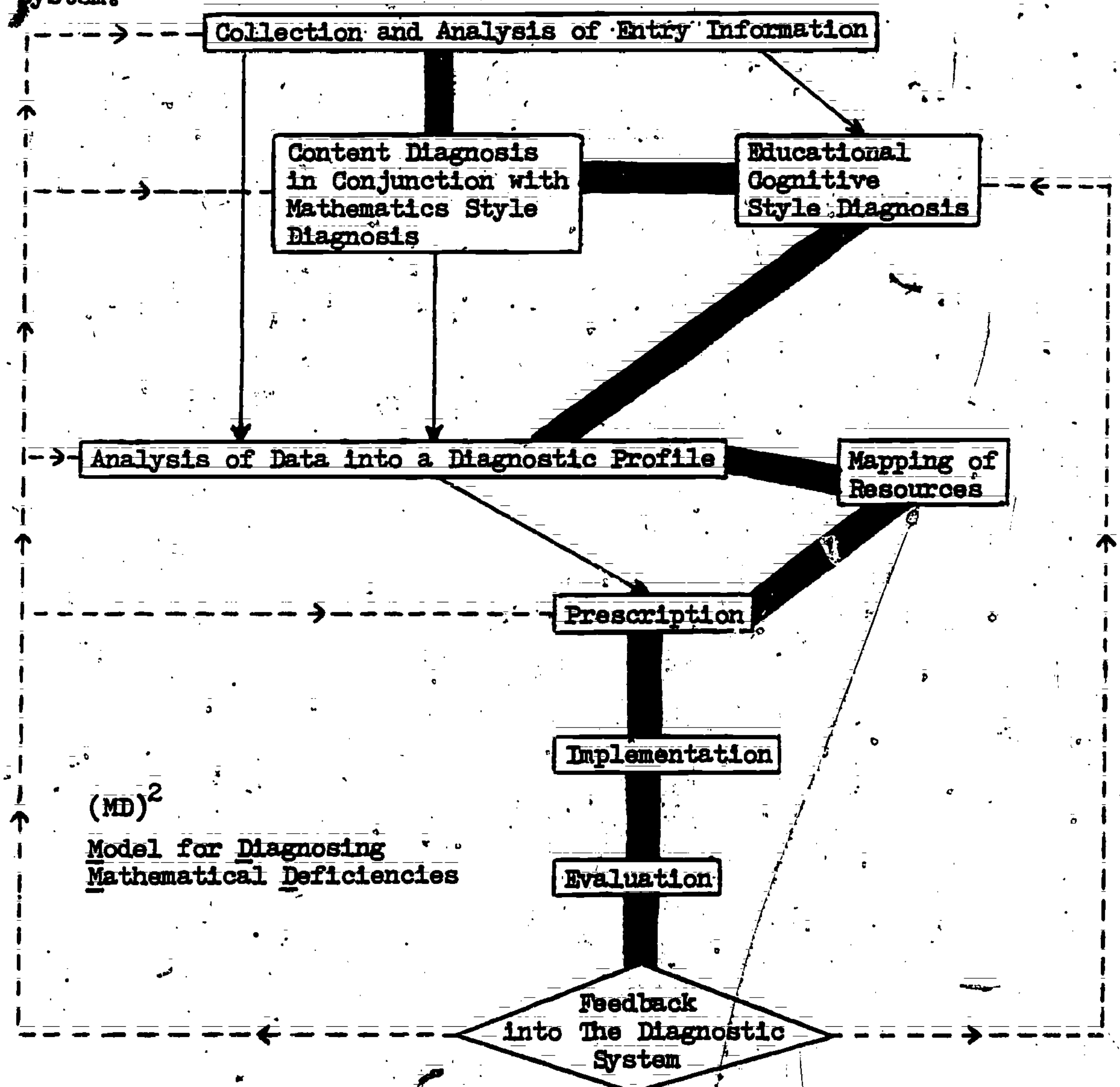
clings to convictions
rejects imposed standards of behavior
prefers to act individually

utilizes both analytical and intuitive processes
finds, organizes, and evaluates relations
has a facility for syllogistic reasoning

The abundance of factors that must be considered in attempting to provide each student with a successful mathematics experience point to the need for a comprehensive diagnostic model at the clinical level (Underhill, 1976).

Some of the questions which must be attended to in the development of such a model include: (1) what general information should be sought, (2) what mathematics content should be considered, (3) what sequence of concepts should be used, (4) what levels of abstraction should be checked, (5) what sensory inputs should be examined, (6) what consideration should be given to affective aspects, (7) what consideration should be given to psycho-motor aspects, and (8) what type of prescriptions should be available. These, and other related questions, form the focusing point for the development of the $(MD)^2$ clinical diagnostic model--A Model for Diagnosing Mathematical Deficiencies.

The model is a three-fold model of diagnosis in that the profile generated for an individual includes data relative to that individual's unique mathematical deficiencies, "mathematics style," and educational cognitive style. This information is gathered and utilized during the course of an intensive analysis of a student using case study techniques. The nine step process involves: (1) collection and analysis of entry information, (2) content diagnosis in conjunction with "mathematics style" diagnosis, (3) educational cognitive style diagnosis, (4) analysis of data from steps 1, 2, and 3 into a diagnostic profile, (5) mapping of resources, (6) prescription, (7) implementation, (8) evaluation, and, if necessary (9) feedback into the diagnostic system.



The first step, the collection and analysis of entry information, is designed to provide general background data concerning the student. Some factors which are included here are parental information, school records, behavior patterns, interests, and anecdotal information such as expressive ability, motivation, self-confidence, attentiveness, and attitude toward mathematics. Every effort should be made to secure reliable data from the student's classroom teacher, school officials, and parents. However, the diagnostician should also engage in the observation of the student with an eye not only toward the analysis of mathematical abilities but also toward those behaviors reflecting the child's physical, psychological, affective, and social orientations. Through an initial interview, as well as other means, the diagnostician should seek to identify such factors as interest, cooperative effort, persistence, flexibility, and adjustment to interview situations. The analysis of the student's reaction to successes, failures, positive reinforcement and negative reinforcement may also provide clues to underlying content deficiencies.

Referral to the $(MD)^2$ model implies that the instigator of the referral has identified some general mathematics content deficiencies or that the instigator simply wants some particular content area diagnosed. These general content areas must be defined and recorded before the $(MD)^2$ model can be implemented. One aspect of this defining process is the interviewing of teachers and parents focusing on their interpretation of what content should be diagnosed. This should then be followed up by an analysis of the student's standardized test results. If no such results are available the diagnostician may request or conduct a standardized diagnostic test such as the Buswell-John Diagnostic Test, KeyMath, or the Stanford Diagnostic Test. This collection of data on the student's content deficiencies is most critical because the analysis of these results provides the means for determining where to begin in step two.

Prior to step two, the student's standardized test results are further analyzed in order to form a profile of generalized mathematics deficiencies which is keyed to the Kent State University Mathematics Checklist (1975) in an attempt to bracket these deficiencies with specific content statements. For example, it may be known that the student has some sort of difficulty with addition involving regrouping. The clinician then translates this information into the relative sections of the Checklist, e.g., place value and addition, and selects appropriate entries which elaborate on the general difficulties, e.g., renaming numerals in several different ways, naming the sum of a two-place whole number and a one-place whole number with single regrouping (ones to tens), and naming the sum of a three-place whole number and a two-place whole number with two regroupings. Thus, the purpose at this point is to tentatively identify those elements of the mathematics checklist which will be used in step two of the model. (It should be noted that the KSU Checklist is a 30 page comprehensive checklist of mathematical concepts for grades K-8.)

Step two represents one of the most critical components of the $(MD)^2$ model for it is at this stage that the specific mathematics deficiencies are isolated. This isolation process occurs through oral interview of the student centering on questions designed to translate checklist entries into specific tasks at the concrete, representative, and abstract levels.

In conjunction with the identification of specific mathematics deficiencies, the clinician identifies the student's "mathematics style." This refers to the observation of the student's utilization of what are referred to as "response modes" and "response formats" in reaction to various "presentation formats."

The $(MD)^2$ model utilizes the following operational definitions for "presentation formats" and "response modes and formats."

Presentation formats can be identified as the following:

Auditory (A): those questions which are posed solely through oral means. The student is asked to respond to that which is heard.

Visual (V): those questions which are posed solely through visual means. The student is asked to respond to that which is seen.

Auditory-Visual (A-V): those questions which are posed through both oral and visual means. The student is asked to respond to that which is heard and seen.

Forced Response: the student must use a given response mode.

Open Response: the student may select a response mode.

Generative: those questions which call for the student to generate the correct response.

Non-generative: those questions which call for the student to select the correct response from a given set of responses.

Response Formats and Modes can be identified as the following:

Oral Concrete (OC): the student responds to a given question by orally describing the situation in terms of concrete objects.

Oral Representative (OR): the student responds to a given question by orally describing the situation in terms of a model or pictorial representation of concrete objects.

Oral Abstract (OAb): the student responds to a given question by orally describing the situation in terms of abstract symbols.

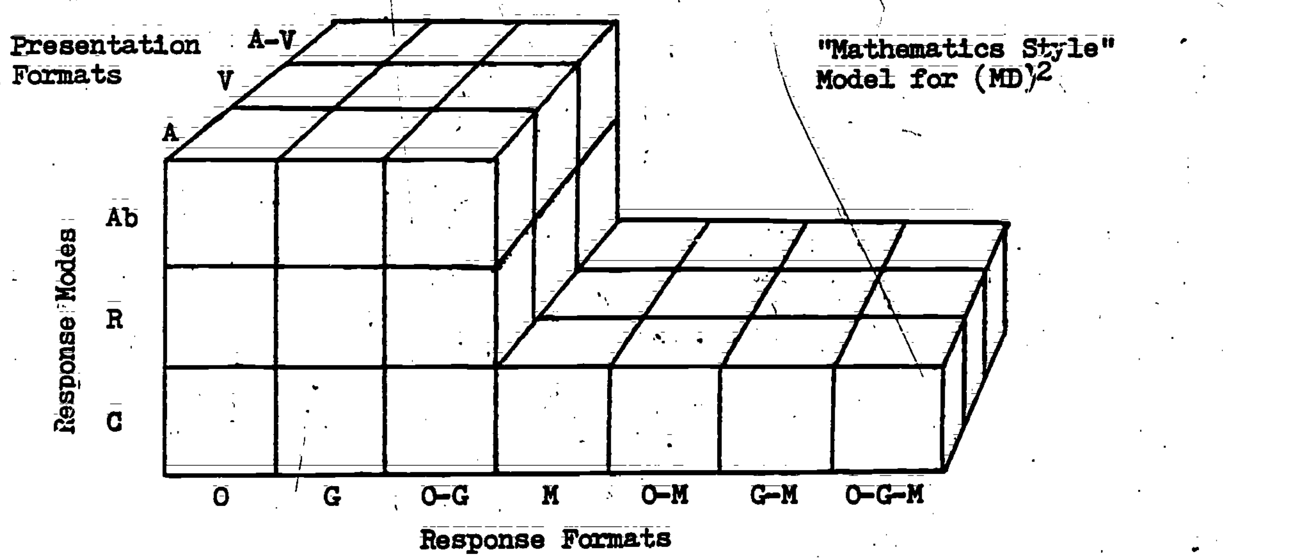
Graphic Concrete (GC): the student responds to a given question by describing, in graphic form, the situation in terms of concrete objects.

Graphic Representative (GR): the student responds to a given question by describing, in graphic form, the situation in terms of a model or pictorial representation of concrete objects.

Graphic Abstract (GAb): the student responds to a given question by describing, in graphic form, the situation in terms of abstract symbols.
Manipulative Concrete (MC): the student responds to a given question by manipulating concrete objects to describe the situation.

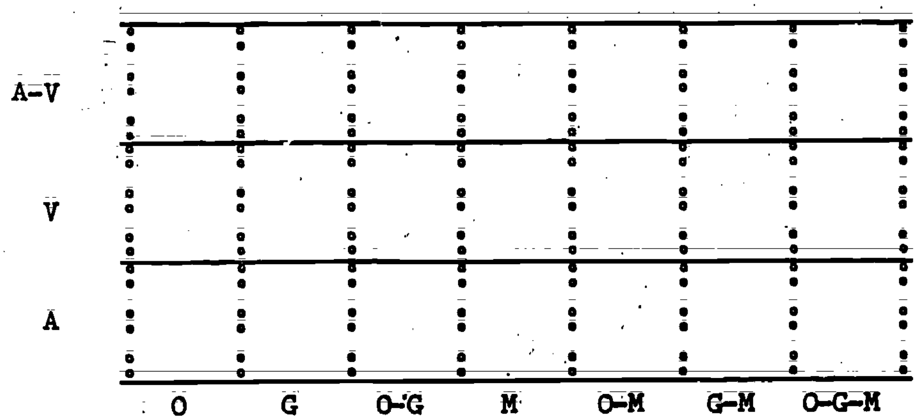
There are several response modes which involve combinations of the aforementioned "unary" response modes. These are:
Oral-Graphic Concrete (O-GC): the student responds to a given question by describing, both orally and in graphic form, the situation in terms of concrete objects.
Oral-Graphic Representative (O-GR): the student responds to a given question by describing, both orally and in graphic form, the situation in terms of a model or pictorial representatio ns of concrete objects.
Oral-Graphic Abstract (O-GAb): the student responds to a given question by describing, both orally and in graphic form, the situation in terms of abstract symbols.
Oral-Manipulative Concrete (O-MC): the student responds to a given question by describing, both orally and by manipulation, the situation in terms of concrete objects.
Graphic-Manipulative Concrete (G-MC): the student responds to a given question by describing, both in graphic form and by manipulation, the situation in terms of concrete objects.
Oral-Graphic-Manipulative Concrete (O-G-MC): the student responds to a given question by describing, in oral and graphic form and by manipulation, the situation in terms of concrete objects.

Graphically, the $(MD)^2$ mathematics style model is shown below.



The previous plate graphically represents the mathematics style component of the $(MD)^2$ model. Note that in a diagnostic session, the trained clinician will use subjective judgment in the selection of which cells to diagnose. For a given checklist entry certain cells are inappropriate and still others can be eliminated on a selective basis. Thus, for a given checklist entry the diagnostician may ask questions based on from one to, say, four cells of the model.

As the content diagnosis progresses the clinician needs to record, in some detail, the events which are (or may be) relevant to identifying the student's deficiencies. The figure below represents one method of recording the presentation format and response mode and format for each major question. It should be noted that the figure has been reduced in size.



Cells are filled with either Ab, R, or C followed by a numeral indicating the order in which the question was asked.

Clearly, the information recorded on this form is invaluable for profiling a student's mathematics style. It does not, however, provide information relative to specific questions asked nor does it make record of any extraneous factors which may affect student response. The Interview Record sheet indicates that the clinician should record the number of the question being asked. Since all of the questions cannot be determined in advance--they will depend on student responses to previous questions--some method of recording questions is needed. For this reason it is strongly suggested that the interview be audio-taped and, if possible, video-taped. This will allow the clinician to reconstruct the session for purposes of further analysis.

When step two of the $(MD)^2$ model is completed the clinician should have a relatively clear picture of the student's mathematical strengths and weaknesses as well as an understanding of the student's "mathematics style." Consequently, it would be possible to terminate the diagnosis at this point and prescribe corrective teaching based on these findings. However, the $(MD)^2$ model has an additional component which enhances the possibility for successful prescriptions--the diagnosis of general educational cognitive style.

Step three, the educational cognitive style component of the $(MD)^2$ model, is based on an abbreviated version of the model proposed by the American Educational Science Association. The clinician should gather cognitive style data through observation and interview whenever possible, however there is a preference test which can be used either solely or in conjunction with the other two methods.

Following the observation, interview, and preference testing of the individual to gather data on educational cognitive style the clinician is prepared to initiate step four of the $(MD)^2$ model. This step represents the stage during which the data collected in steps one, two, and three is analyzed into a student diagnostic profile. This profile includes data pertaining to: (1) general information reflecting student background, (2) the student's specific mathematics deficiencies and mathematics style, and (3) the student's educational cognitive style. The clinician's task is to piece together this information to form a profile representing the diagnosis of the individual.

One aspect of stage four is the search for consistency between the student's mathematics style and educational cognitive style. A given student may, during the content diagnosis, exhibit a tendency to react positively to questions presented in a visual format but negatively to those with an auditory format. If this same student's educational cognitive style map indicates a minor or negligible T(VL), T(VQ), or Q(V) element then the diagnosis may be incomplete--at the least, it must be reviewed for errors. If, on the other hand, the findings from the mathematics style diagnosis and the findings from the educational cognitive style diagnosis match then the clinician can be reasonably certain that steps two and three of the $(MD)^2$ model were successful.

An additional element of consistency can be checked at this stage of the diagnostic process. It is possible to describe a cognitive style map which indicates the ability to deal with mathematics presented at the concrete, representative, and abstract levels. Using the information provided by such maps the clinician is able to determine whether an individual's inability to successfully deal with mathematics presented at a particular level of abstraction is caused by a deficiency in certain cognitive style components or by a lack of experience with a given level of abstraction. That is to say, if an individual's map indicates the presence of those components necessary for working at the concrete level but the mathematics style aspect of the content diagnosis suggests that the student has difficulty describing or completing exercises at this level then, perhaps the student has not had sufficient experiences with concrete models. The concept of division serves as an excellent example of the importance of this step of the analysis done by the clinician. Consider a student that can complete the division algorithm but cannot use counters to illustrate the process. Is it because the student is incapable of modeling (because of a lack of cognitive style components) or is it because the student has never had to model the operation and therefore lacks understanding and experience? The question poses interesting problems in designing corrective procedures.

Step five of the $(MD)^2$ model represents the first step in the process of prescription development. Diagnosis based on cognitive style constructs will be of little value if prescribed activities do not provide a high probability of student success. Therefore, the diagnostician must not only diagnose the student, but also those tasks which may be used in subsequent instruction (Hill, 1974). In this manner cognitive style diagnosis not only provides for the identification of the unique structures each individual brings to a learning situation, but it also allows for the translation of this uniqueness into proposed programs of instruction (Radike, 1973).

In step five, the clinician maps the instructional resources--the purpose being to determine the cognitive style conditions of those aspects of the educational environment which may be used for corrective teaching. Included in this category, and therefore subject to mapping, are persons (those individual's that may play a role in the subsequent teaching of the child, e.g., teacher, tutor, librarian, and counselor), processes (those activities which may be used in subsequent teaching, e.g., methods of instruction), and properties (materials used in subsequent teaching, e.g., audio-tapes, films, books, worksheets, and manipulatives). Generally speaking, diagnosis of properties provides data for the matching of symbolic orientations, diagnosis of processes provides data for matching modalities of inference, and diagnosis of persons provides data for matching cultural determinants. It should be noted, however, that these three components of cognitive style must be considered as inseparable and consequently need to be viewed as a totality.

The mapping process at this stage is the same as it was for step three. The clinician must map the instructional resources in the same manner as mapping the student who will come into contact with these resources. That is, the symbolic conditions of elements of the instructional process are determined so that individuals can be matched to these for prescriptive purposes. As an example, assume a possible corrective teaching technique involves having the child work with a peer on the analysis of word problems according to certain delineated procedures--a fixed step-by-step approach. The clinician mapping such a task may arrive at the following condition of this task.

$$\begin{Bmatrix} T(VL) \\ T(VQ) \\ Q(V) \\ Q(CET) \end{Bmatrix} \times \begin{Bmatrix} A \\ F' \\ \\ \end{Bmatrix} \times \begin{Bmatrix} M \\ \\ \\ \end{Bmatrix}$$

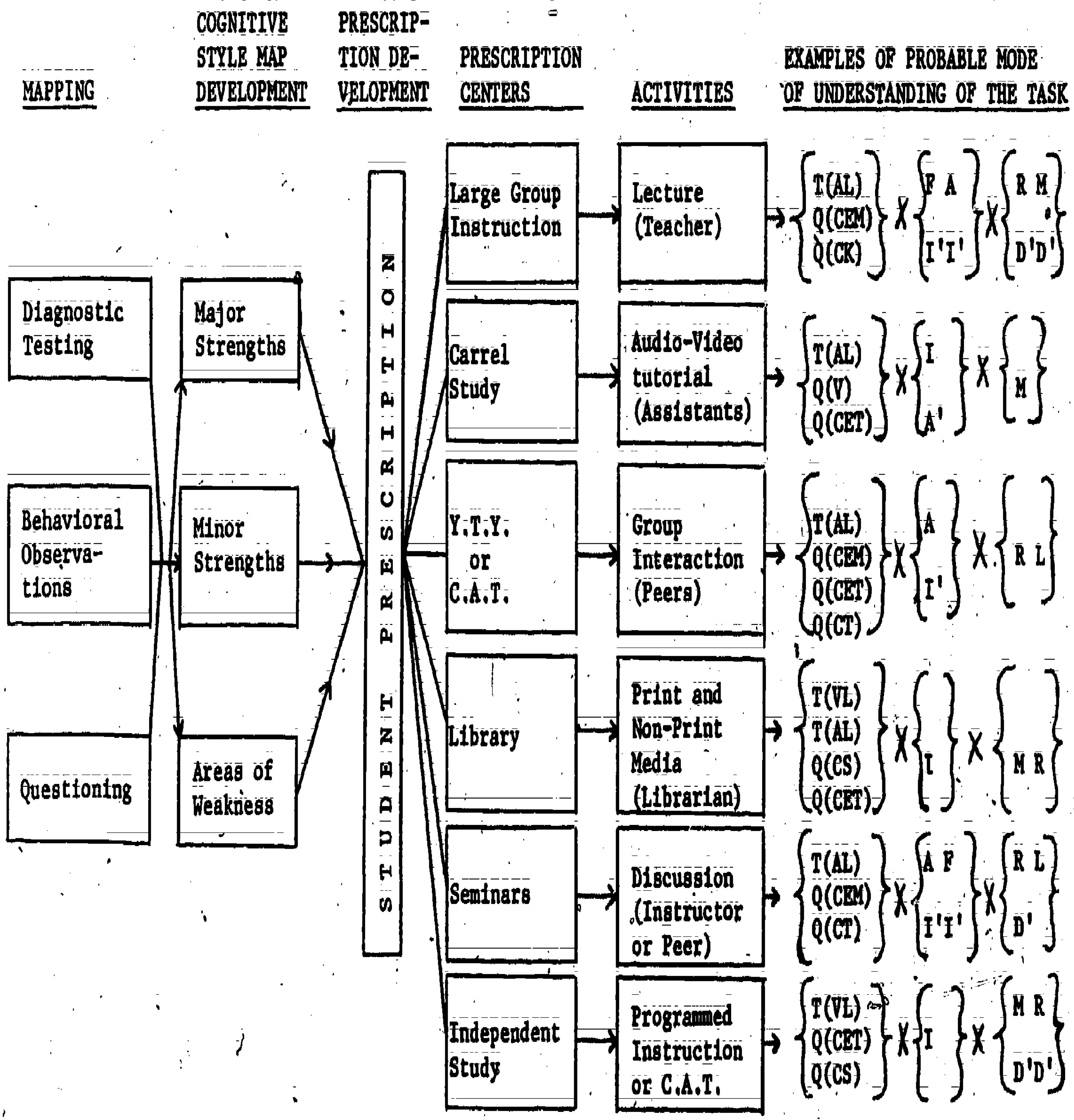
For the child whose map is shown below this may not be a valuable experience.

$$\begin{Bmatrix} T(AQ) \\ Q(A) \\ Q(T) \end{Bmatrix} \times \begin{Bmatrix} I \\ \\ \end{Bmatrix} \times \begin{Bmatrix} R \\ \\ \end{Bmatrix}$$

That is, the interface between the conditions of the task and the cognitive style of the student do not sufficiently match.

It should be noted that, theoretically at least, step five may only have to be completed once. If the clinician, or the teacher, maps all of the resources available then these same maps can be filed for use in the development of prescriptions for several students. Therefore, a goal of the clinician should be the compilation of resource maps for quick reference in future situations.

Step six of the $(MD)^2$ model involves the preparation of prescriptions and, of course, these prescriptions are based on the information gleaned from steps one through five. Cognizance of individual differences is no less critical at this stage than it is at any other stage of the $(MD)^2$ model. The purpose of the prescription is to relate the individual's unique $(MD)^2$ profile to the available resources in a manner which will provide the greatest probability of success. Some possible prescriptive routes based on cognitive style diagnosis of both individuals and educational tasks are presented in the following table.



Personalizing Education (Adapted from Bowman, Birch, Hill, & Nunney, 1974)

The format of the prescription may vary but there are certain characteristics which each prescription should include. One criteria is comprehensiveness. The prescription should include a synopsis of all data obtained in steps one through five organized and presented in a manner which clearly conveys the clinicians analysis and interpretation of the data. Any formal and/or informal tests which were administered during the course of the diagnosis should be described along with student reactions to these tests. Each major component of the $(MD)^2$ model, i.e., content diagnosis, mathematics style, and educational cognitive style, must be discussed in great detail. Each component could be mentioned in isolation, however, because of their interrelatedness, a concurrent discussion designed at weaving the three into one overall picture is preferred. Separate discussions of the components does, however, have the advantage of providing the teacher with a clear-cut picture of each. Also, because of the strong relationship between mathematics style and educational cognitive style it may be desirable to present a cogent discussion of these two components and a separate discussion of specific content deficiencies. Nevertheless, this should still be followed by a synthesis of all three as a hedge against the teacher basing corrective teaching on the content diagnosis alone. The teacher must not overlook the prescriptive benefits gained from the model's ability to diagnose the student's capabilities in the areas of symbolic orientations, modalities of inference, and large group, small group, and independent study.

One final point on prescription deserves--indeed demands--attention. Any diagnosis will indicate to the clinician certain strengths and weaknesses which will necessarily affect the design of corrective teaching procedures--the question is "How"? The clinician has two alternatives to pursue.

On the one hand, a prescription can be written which calls for corrective teaching techniques designed to utilize strengths in both mathematics style and cognitive style to build skill and understanding with those areas identified as content deficiencies, for the most part ignoring style weaknesses in the instructional design. For example, consider the student with the following over-simplified profile.

Content Deficiency
- Renaming fractions in simplest form
- Comparing fractional numbers

Mathematical style
- Auditorily oriented
- Uses concrete objects

Educational Cognitive Style

$$\left\{ \begin{matrix} T(AL) & T'(VQ) \\ Q(A) & \\ Q(CET) & \end{matrix} \right\} \quad X \quad \left\{ \begin{matrix} I \\ \\ \end{matrix} \right\} \quad X \quad \left\{ \begin{matrix} M \\ \\ \end{matrix} \right\}$$

Under the above philosophy of corrective teaching this student may be asked to sit alone at a listening station with audio tapes keyed to cards with appropriate numerals on them. These tapes describe the process of renaming fractions in simplest form by way of using concrete objects and directs the child to model fractions using cuisenaire rods placed at the station. Little or no reference would be made to visually-oriented resources, e.g., filmstrips, nor would any concerted effort be made to have the student work with peers or watch the teacher model some examples at the board.

The other alternative position on corrective teaching is to utilize both strengths and weaknesses in mathematics style and cognitive style in order to eliminate content deficiencies. For the example just given the prescription would include such activities as peer assistance, visual aids, teacher demonstration, etc. The rationale for using style weaknesses is that through use they may develop into strengths.

On the surface it seems as though the second alternative would be most beneficial. In fact, in most instances it would be the most probable route for eliminating deficiencies. After all, it does provide the student with a greater variety of opportunities to identify errors and misconceptions. However, this does not mean that this philosophy of corrective teaching will work best for all. Those students with serious content deficiencies may become even more confused by having to deal with two deficiencies at once, namely, content deficiencies and style deficiencies.

To summarize, prescriptions calling for corrective teaching demand careful consideration. At the risk of over-simplifying, students with major content deficiencies should receive corrective instruction designed to utilize their style strengths to alleviate these deficiencies, while students with minor content deficiencies should receive corrective instruction designed to utilize their style strengths and weaknesses to alleviate their deficiencies. The determination of what are major or minor content deficiencies should be based on analysis of test results and the subjective judgment of a trained clinician.

Step seven of the $(MD)^2$ model, implementation, refers to the means of incorporating the results of the entire diagnostic process into the instructional program. This is not to be confused with the act of corrective teaching which is not a part of the model since it is an activity carried out by the classroom teacher. The implementation of the $(MD)^2$ model refers to the manner in which the diagnosis is transmitted to the teacher. The vehicle for accomplishing this step is the clinician-teacher conference. The prescription report discussed in step six may simply be delivered to the teacher for consideration, however, it is strongly suggested that a conference be arranged so that the possibilities of misinterpretation are lessened prior to the initiation of corrective teaching procedures. When both parties have come to a consensus concerning the findings of the diagnosis and the purpose and rationale of the prescription the teacher may then determine the manner in which instruction will take place.

Evaluation, step eight, is based primarily on teacher-input after corrective teaching has begun, once again through clinician-teacher conferences. The teacher input should be founded on observations and test results. An additional aspect of the evaluation step is observation, by the clinician, of the child at work in the classroom. Observation by both the clinician and the teacher is designed to allow for a "comparing of notes" (which may assist the development of greater interrater reliability for later referals) and to form a common base of knowledge concerning the student's status to insure the success of the clinician-teacher conference.

Feedback into the diagnostic system, step nine, may be the result of step eight, the clinician-teacher conference. The re-entry step will vary for individual students. For some it may be necessary to begin the entire diagnostic process anew. For others, re-entry may take place at either step two (content diagnosis incorporating mathematics style), step three (educational cognitive style diagnosis), step four (analysis of steps one, two, or three), or step six (prescription writing).

Thus, the diagnostic process has traveled full circle. If carried out properly there should be a wealth of information available concerning not only what the student does and does not know but also concerning the manner in which the student does and does not take on meaning. The real value, however, is not in simply having this information but in using it. A carefully planned follow-up program of corrective teaching is critical to the success of any diagnostic venture.

<u>Conclusion and Recommendations for Further Study</u>

By way of conclusion it seems appropriate to review some of the more pertinent aspects of the model. First, it is essential that the reader understand the clinical nature of the model. It is designed to describe a possible diagnostic process which can be carried out by a trained clinician--it may have some value in classroom diagnosis but this is not the intended target. Second, it is important to note that this model is not intended to be used only with those students that have severe mathematics difficulties. The mathematics style and educational cognitive style components of the model make this model valuable for the diagnosis of any student of mathematics. Third, in the actual diagnosis of an individual's mathematics style it is not intended that all thirty-nine cells be tested. It is necessary for the clinician to use subjective judgment and select those cells which are most appropriate for the task at hand. Fourth, the $(MD)^2$ model, when implemented, does not describe a diagnostic test from which one teaches. There must be intermediate steps between the administration of this model and the actual instructional process.

The Model for Diagnosing Mathematical Deficiencies is presented as one conceptualization of the manner in which diagnosis could proceed. The evidence on interrelationships among the abundance of factors affecting development suggests that each student brings to each learning situation a differential combination of unique capabilities and abilities, each at a particular stage of development. Diagnosis, then, should strive to describe these capabilities and abilities and the factors which affect them. Its ultimate purpose is to facilitate the construction of individualized prescriptions for uniquely organized persons. The $(MD)^2$ model represents one attempt to reach this goal.

The purpose of this paper was to develop and describe a clinical model for diagnosing mathematical deficiencies which incorporates cognitive, affective, and psychomotor aspects of educational cognitive style. This model is designed in such a way as to reflect consistency with the view that the task of diagnosis is to describe a personality as well as a person's subject matter deficiencies.

It is important to note that the purpose of this paper was not to describe processes through which the diagnostic model could be implemented. Indeed, application concerns are not (nor should they be), factors in the design stage of model development. Hypothesizing on possible application difficulties prior to the development of a model may cause undue restrictions and limitations to form in the mind of the designer. Comments relevant to this point can be found throughout the literature. For example,

> A basic innovative design may well be 'useless' in the sense that it has little or no application immediately to schools as educational institutions.... Concern for the immediate applicability of the findings can distract the researcher, narrow his efforts and hasten him to unjustified conclusions. (Brickell, 1961, p. 82)

Thus, Brickell described three distinguishable phases of innovation: design, evaluation, and demonstration; and their ideal settings which are, respectively: freedom, control, and normality.

Werner (1968), in commenting on Brickell's notion of three phases of innovation, stated

> Design efforts cannot be conducted in evaluation settings because experimental controls of the type needed for adequate evaluation are restrictive by their very nature. These restrictions reduce the freedom to explore for something better. The ordinary, unenriched setting needed for the demonstration of a proven innovation is the setting least likely to generate new designs. The observer of a demonstration needs to see the demonstration of the innovation as part of the normal, ongoing program in a school like his own. For these reasons, therefore, the circumstances needed for the design of an innovation cannot be reconciled with those needed for proper evaluation and demonstration of the innovation. (Werner, 1968, pp. 89-90)

A similar concern was expressed by Wilson (1973) in a paper calling for more efforts in generative research--specifically clinical intervention research. Citing the unfortunate tendency to evaluate the results of generative research on the basis of criteria designed for confirmative experimental research, Wilson noted that "clinical intervention is either not recognized as research, or is depreciated as research, and in some circles even deprecated as research" (Wilson, 1973, pp. 1-2).

As previously mentioned, Engelhardt (1974) has commented on the difficulties encountered in experimental research without the benefit of a theoretical model. Thus the significance of this paper lies in its purpose, that is, it is the development of a theoretical model, and a description of its accompanying instrumentation, for clinically diagnosing mathematical deficiencies and its relation to the teaching and to the learning of mathematics.

Thus, at the risk of "depreciation" or "deprecation," the present paper is best described as a generative effort to describe a diagnostic model for clinical use in identifying an individual's unique mathematics deficiencies. It should be recognized as a first attempt which has, at this point, only been administered on a limited basis. Now, and only now, it needs to be examined, and possibly adapted, for use in diagnostic-prescriptive settings.

Because this study is not based on statistical analysis conclusions similar to those found in experimentally-oriented studies cannot be stated. Consequently, this section is devoid of any attempt to state inferences but instead focuses on areas of further study and needed research.

The design process used in the development of the $(MD)^2$ model suggests the following areas of needed research and development.

1. The $(MD)^2$ model suggests a hierarchical checklist for use in determining a student's content deficiencies. This hierarchy is based on expert judgment but perhaps other means of hierarchy validation such as Guttman analysis should be attempted.

2. The $(MD)^2$ model suggests the use of a standardized diagnostic test to obtain entry level information on the child's content deficiencies. Does a standardized diagnostic test provide sufficient data for transferring general difficulty areas into the checklist?

3. The $(MD)^2$ model suggests instrumentation to be used in conjunction with a checklist for diagnosing a student's mathematical style. A next step would be the development of a battery of items for each entry in the checklist.

   Several questions would need attention:

   (a) Should the development of this battery begin with one or two concept clusters or should the entire checklist be subject to item development? (This question becomes critical when one realizes that the checklist is not sequenced across concepts.)

   (b) How many items are needed for each entry of the checklist? (It is necessary to consider the possible presentation and response formats for an individual entry before this question can be answered.)

   (c) Concerning the internal structure of the item battery, should the questions be open or forced response? Generative or non-generative response? Should there be a mixture of these response types?

Considering implemention of the $(MD)^2$ model in diagnostic prescriptive settings, the following areas need to be inspected.

1. The $(MD)^2$ model, while not so complicated that it can't be implemented, does require a certain amount of expertise. What procedures must be developed for training mathematics clinicians in the use of the $(MD)^2$ model? For training classroom teachers in its use? In what ways can computer capabilities be used to simplify the data collection and record keeping aspects of the $(MD)^2$ model?

2. The $(MD)^2$ model is designed for use in a mathematics clinic, or at least in a program with clinical procedures. Several authors have suggested that the adaptation of clinical practices to classroom techniques is beneficial (Buswell, 1935; Callahan, 1973; Denmark, 1974). What adaptation is necessary before $(MD)^2$ can be used in classroom diagnosis?

3. The educational cognitive style component of the $(MD)^2$ model has been successfully used at the upper elementary, secondary, and college levels. The $(MD)^2$ model itself is designed for use at any level of mathematics instruction, however the vehicle used to describe the model is elementary school mathematics. What changes, if any, are needed in the model design before it can be implemented at the secondary and post-secondary levels?

4. It has been suggested that longitudinal study is needed to determine the usefulness of cognitive style diagnosis (Sternberg, 1975). Thus, a long-range testing program may need to be established before the value of a model such as $(MD)^2$ can be fully evaluated.

Assuming that implementation attempts are successful, certain questions on the actual use of the $(MD)^2$ model need to be addressed.

These include:

1. The $(MD)^2$ model is designed to provide an individual profile of three major areas: content deficiencies, mathematics style, and educational cognitive style. Is there benefit to be gained by fractionating the model and using only one of the components? Any two of the components?

2. The $(MD)^2$ diagnostic model assumes that, among other methods, a dyadic interview will be used to gather data. If a student does not perform appropriately on a diagnostic instrument it may be due to other factors aside from the student not understanding the concept being assessed. What critical factors can be identified that affect clinician-student interactions? What role does clinician cognitive style play in determining the success of the interview? Is there a "most effective" clinician cognitive style?

3. The $(MD)^2$ model is designed to describe individuals from a variety of perspectives. What are the individual difference variables which might affect performance in the diagnostic setting that have not been considered? How does use of the model affect children? Are cognitive styles of individuals content specific?

4. The $(MD)^2$ model provides the opportunity to use diagnosed strengths and weaknesses in a variety of ways. Does remediation based on using cognitive style strengths to eliminate content deficiencies prove to be more successful than using a combination which stresses strengths but also attempts to eliminate weaknesses through use?

The above represent questions generated from the development of the Clinical Model for Diagnosing Mathematical Deficiencies--$(MD)^2$. Since this paper is concerned with only the design stage of model development it is clear that a significant amount of work must be done before implementation. This work now begins.

REFERENCES

AESA Membership Directory. Bloomfield Hills, Mi.: American Educational Sciences Association, October, 1975.

Apple, M. The adequacy of systems management procedures in education. In A. Yee (Ed.), Perspectives on management systems approaches in education. Englewood Cliffs, N.J.: Educational Technology Publications, 1973.

Ashlock, R. B. Structures for teacher education related to diagnosis and remediation: Needed research. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Berry, J. J., Sutton, T. J., & McBeth, L. S. (Eds.). The educational sciences: A bibliography with commentary (Rev. ed.). Bloomfield Hills, Mi.: American Educational Sciences Association, 1975.

Bowman, B. A., Birch, B. A., Hill, J. E., & Nunney, D. N. Personalized educational programs. East Lansing, Mi.: Teacher Improvement Project, 1974.

Brickell, H. M. Organizing New York state for educational change. Albany, N.Y.: State Department of Education, 1961. Cited by Werner, M. J., An application of critical path analysis to the design of a systematically articulated curriculum in science and mathematics for secondary schools (Doctoral dissertation, Boston College, 1968). Dissertation Abstracts, 1969, 29, 4209A. (University Microfilms No. 69-09833)

Burns, H. W. The logic of educational implications. Educational Theory, 1962, 12, 53-63.

Buswell, G. T. The place of the psychological laboratory in educational diagnosis. In G. M. Whipple (Ed.), Educational diagnosis. The thirty-fourth yearbook of the National Society for the Study of Education. Bloomington, IL.: Public School Publishing, 1935.

Callahan, L. G. Clinical evaluation and the classroom teacher. Paper presented at the meeting of the Educational Research Association, New Orleans, Louisiana, February, 1973.

Callahan, L. G. Conference summary. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, May, 1974.

Callahan, L. G., & Robinson, M. L. Task-analysis procedures in mathematics instruction of achievers and underachievers. School Science and Mathematics, 1973, 73, 578-584.

Capps, K. R. Thoughts on coordinating a research effort in remediation in mathematics. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Cawley, J. F. Learning disabilities in mathematics: A curriculum design for upper grades. A proposal submitted to the Office of Education for the Handicapped, Washington, D.C., October, 1975.

Clements, M. Theory and education. Educational Theory, 1962, 12, 124-128.

Commission on Preservice Education of Teachers of Mathematics of the National Council of Teachers of Mathematics. Guidelines for the preparation of teachers of mathematics. Reston, Va.: National Council of Teachers of Mathematics, 1973.

Crenson, J. Diagnostic prescriptive mathematics teaching. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Cronbach, L. J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Cunningham, W. G. The impact of student-teacher pairings on teacher effectiveness. American Educational Research Journal, 1975, 12, 169-189.

Denmark, T. Classroom diagnoses: A reaction paper. A paper presented at the Conference on Remedial Mathematics, Kent, Ohio, May, 1974.

Engelhardt, J. M. Remediation of learning difficulties in school mathematics: Promising procedures and directions. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, May, 1974.

Engelhardt, J. M. Diagnosis and remediation in school mathematics: Developing continuity among r and d efforts. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Florida Statutes Section 230.2317. As amended by Chapter 75-284, Laws of Florida, paragraphs 3 and 4, 1975.

Golladay, M. A., Devault, M. V., Fox, T., Jr., & Skuldt, K. Problems in empirical research on individualized mathematics programs. Journal for Research in Mathematics Education, 1975, 6, 159-169.

Gowin, D. B. Can educational theory guide practice? Educational Theory, 1963, 13, 6-12.

Gowin, D. B. Defects in the doctrine of educational strategies. Educational Theory, 1964, 14, 95-98.

Gray, J. F. Issues in mathematics education. Paper presented at the meeting of the National Council of Teachers of Mathematics Forum on Teacher Education, Chicago, Illinois, April, 1972.

Guttchen, R. S. The quest for necessity. Educational Theory, 1966, 16, 128-134.

Hancock, R. R. Cognitive factors and their interaction with instructional mode. Journal for Research in Mathematics Education, 1975, 6, 37-50.

Heddens, J. W. Clinical diagnosis of children with mathematics difficulties. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, May, 1974.

Heddens, J. W. Personal interview, Kent State University, Kent, Ohio, December 5, 1975.

Hill, J. E. An outline of the educational sciences. A proposed conceptual framework for education. Bloomfield Hills, Mi.: Oakland Community College Press, 1968.

Hill, J. E. Cognitive style as an educational science. Bloomfield Hills, Mi.: Oakland Community College Press, 1970.

Hill, J. E. The educational sciences. Bloomfield Hills, Mi.: Oakland Community College Press, 1972.

Hill, J. E. Reliability, validity, objectivity, and discriminative factors associated with the process of mapping "styles." Bloomfield Hills, Mi.: Oakland Community College Press, 1973.

Hill, J. E. The educational sciences. Bloomfield Hills, Mi.: Oakland Community College Press, 1974.

Hill, J. E. Personal interview, Oakland Community College, Bloomfield Hills, Michigan, March 8, 1976.

Hill, W. F. Learning--A survey of psychological interpretation. San Francisco: Chandler Publishing, 1963.

Hynes, M. C. Response modes and diagnostic procedures. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Irons, J. L. The need for packaged material and a delivery system. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Johnson, M. L. The role of Piagetian theory and 'Piagetian-type' tasks in mathematics diagnosis and remediation. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Kent State University Mathematics Education Team. Checklist of mathematical concepts. Unpublished manuscript, 1975. (Available from Kent State University, Kent, Ohio).

Krech, D. Psychoneurobiochemeducation. Phi Delta Kappan, 1959, 50, 370-375.

Mouly, G. J. Psychology for effective teaching (2nd ed.). Atlanta: Holt, Rinehart & Winston, 1968.

National Advisory Committee on Mathematical Education. Overview and analysis of school mathematics. Washington, D.C.: Conference Board of the Mathematical Sciences, 1975.

National Assessment of Educational Progress. Math fundamentals: Selected results from the first national assessment of mathematics. The Education Commission of the States, 1973.

National Conference on Needed Research in Mathematics Education. Journal of Research and Development in Education, 1967, 1, 1-142.

National Conference on Remedial Mathematics. Collection of unpublished papers presented at the First National Conference on Remedial Mathematics, Kent, Ohio, May, 1974.

National Conference on Remedial Mathematics. Collection of unpublished papers presented at the Second National Conference on Remedial Mathematics, Kent, Ohio, April, 1975.

National Conference on Remedial Mathematics. Collection of unpublished papers presented at the Third National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Newsome, G. L. In what sense is theory a guide to practice in education? Educational Theory, 1964, 14, 31-39.

Nunnally, J. C. Psychometric theory--25 years ago and now. Educational Researcher, 1975, 4, 7-13; 19-20.

Perkinson, J. J. A note on: Can educational theory guide practice? Educational Theory, 1964, 14, 90-94.

Piaget, J. The origins of intelligence in children. Translated by Margaret Cook. New York: W. W. Norton & Co., 1952.

Radike, F. W. (Ed.). Handbook for teacher improvement utilizing the educational sciences. East Lansing, Mi.: American Educational Sciences Association, June, 1973.

Reisman, F. K. A guide to the diagnostic teaching of arithmetic. Columbus, Ohio: Charles E. Merrill, 1972.

Reys, R. E., & Post, T. R. The mathematics laboratory: Theory to practice. Boston: Prindle, Weber & Schmidt, 1973.

Riedesel, C. A. Reaction paper to: Clinical diagnosis of children with mathematics difficulty. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, May, 1974.

Romberg, T. A. The application of "process evaluation" to corrective instruction: A suggestion. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Scandura, J. M. An alternative approach to research on mathematics learning. Paper presented at the meeting of the National Conference on Needed Research in Mathematics Education, Journal of Research and Development in Education, 1967, 1, 33-43.

Shuert, K. L. A study to determine whether a selected type of cognitive style predisposes one to do well in mathematics (Doctoral dissertation, Wayne State University, 1970). Dissertation Abstracts International, 1971, 31, 3352A. (University Microfilms No. 71-00454)

Smiler, H. B. A survey of research in mathematics education. In A history of mathematics education in the United States and Canada, appendix b. The thirty-second yearbook of the National Council of Teachers of Mathematics. Washington, D.C.: National Council of Teachers of Mathematics, 1970.

Soltis, J. F. An introduction to the analysis of educational concepts. Reading, Mass.: Addison-Wesley, 1968.

Steffe, L. P., & Smock, C. D. On a model for learning and teaching mathematics. In L. P. Steffe (Ed.), Research on mathematical thinking of young children. Reston, Va.: National Council of Teachers of Mathematics, 1975.

Sternberg, L. Personal correspondence. Northern Illinois University, DeKalb, Illinois, November 13, 1975.

Stiglmeier, L. M. Teacher's judgments of pupil's dependence/self-reliance characteristics mode of instruction and their relationship to achievement (Doctoral dissertation, State University of New York at Albany, 1972). Dissertation Abstracts International, 1973, 34, 1808A. (University Microfilms No. 73-19707)

Suydam, M. N. The status of research on elementary school mathematics. In R. B. Ashlock & W. L. Herman, Jr. (Eds.), Current research in elementary school mathematics. London: The Macmillan Co., Collier-Macmillan Ltd., 1970.

Trueblood, C. R. A model for using diagnosis in individualizing mathematics instruction in the elementary school. The Arithmetic Teacher, 1971, 18, 505-511.

Underhill, R. G. A diagnosis position paper: General needs and hierarchies. Paper presented at the National Conference on Remedial Mathematics, Kent, Ohio, April, 1976.

Uprichard, A. E., Baker, B. E., Dinkel, N. R., & Archer, G. A task-process integration model for diagnosis and prescription in mathematics. Unpublished manuscript, 1975. (Available from the University of South Florida).

Van Dalen, D. B. Understanding educational research. New York: McGraw-Hill, 1966.

Weaver, J. F. Some concerns about the application of Piaget's theory and research to mathematical learning and instruction. The Arithmetic Teacher, 1972, 19, 263-270.

Werner, M. J. An application of critical path analysis to the design of a systematically articulated curriculum in science and mathematics for secondary schools (Doctoral dissertation, Boston College, 1968). Dissertation Abstracts, 1969, 29, 4209A. (University Microfilms No. 69-09653)

Wilson, J. W. An epistemologically based system for classifying research and the role of clinical intervention research in that system. Paper presented at the meeting of the American Educational Research Association, New Orleans, February, 1973.

Wittrock, M. C. A generative model of mathematics learning. Journal for Research in Mathematics Education, 1974, 5, 181-196.

Wohlwill, J. F. Piaget's system as a source of empirical research. In I. E. Sigel & F. H. Hooper (Eds.), Logical thinking in children: Research based on Piaget's thinking. New York: Holt, Rinehart & Winston, 1968.

Reaction Paper
to
A Clinical Model for Diagnosing Mathematical Deficiencies Incorporating Educational Cognitive Style

Lelon R. Capps
University of Kansas

This paper is brief for two reasons. First, the description of the model is indeed complete and comprehensive. Second, much of the information needed for revision and reaction can be forthcoming only as the model is implemented. The last six pages of the paper presented by Dr. Speer outline many of the unanswered questions concerning the application of the model and deserve the careful attention of anyone involved in using the model.

One must note that the model relies heavily on subjective interpretation of verbal description. Thus, some specific training is necessary to gain consistency in descriptors of the behavior. Also, in some instances there may be the possibility of gathering similar data with more objective checklists.

In attempting to describe educational cognitive style the search must be for some pattern of consistency. It is on this component of the model where the greatest care must be exercised in interpreting the information gathered in the process. To predict behavior of this nature with consistency is indeed a difficult task.

While educational cognitive style is content free, it is not culture free. Perhaps some future research should deal with the role of cultural background and its effect on cognitive style. Stodolsky and Lesser (1967) suggest the nature of this relationship in their studies.

In gathering information about content deficiencies, using the standardized tests may be of only limited value. It should be viewed only as a broad screening device. As indicated in the paper, one needs a much more comprehensive set of descriptors such as those included in the Kent State University Mathematics Checklist.

It should be pointed out that current practice among educators usually goes only through step two of this model. It is step three of this model that makes it intriguing and appealing to the researcher and teacher interested in diagnostic and prescriptive teaching.

In step four of the model one wonders if a learner will exhibit the same profile across content areas and across diagnosticians. This will need to be clearly demonstrated by careful collection of data when the model is implemented.

In step six of the model, the diagnosticians must not be misled into thinking that there is one best prescription. In practice the model will allow for a variation of prescription. Of basic interest should be a study of the relationship of the profile to the prescription and the success of the prescription. The reactor suggests that careful research should be conducted to study this relationship.

A major concern in implementing the prescription centers on who does the prescriptive instruction. If the diagnostician does the corrective instruction, the probability for success will be greater. If the classroom teacher is to do the corrective teaching, there is a definite risk of losing much of the information gathered on the profile as it relates to the instructional process. For this reason, transmitting the information to a classroom teacher is a critical step in the process and if not well done, could result in very limited success. Hopefully, the model will be of value to the classroom teacher even though it is a clinical model. Perhaps, the model will be of <u>most</u> value <u>only</u> if the classroom teacher can interpret the results and implement the corrective instruction successfully.

In the end, the model also must be of value in identifying and perhaps modifying the educational cognitive style of the teacher. That is, weakness in instructional strategies being employed by a teacher can be identified and, if modified appropriately, there will be a reduction in the number of students with mathematical deficiencies in the teacher's classes.

In summary, the model presents an initial and hopeful step in the area of diagnosis and prescription. Dr. Speer has done a commendable piece of work though there is much yet to be done. The model makes no claims to be an instrument of an exact science. To the contrary, it is a map which shows the route to follow but does not guarantee the condition of the road. The judgment of the user and the ability to remain flexible no doubt will be two crucial variables in determining the condition of the road.

REFERENCE

Stodolsky, S. & Lesser, G. Learning patterns in the disadvantaged. <u>Harvard Educational Review</u>, 1967, 37, 546-93.

Feedback in Diagnostic Testing

Jon M. Engelhardt
Arizona State University

In the testing situation, telling students whether or not their answers are correct is commonly prohibited or delayed until students have responded to all the items (e.g., Anastasi, 1976; Cronbach, 1960). Although it is sometimes suggested that testers insert perfunctory comments like "Good" or "Fine," this is done only to encourage the students by acknowledging effort, not to indicate success on a particular item. The primary rationale for withholding such correctness feedback is to "standardize" testing conditions, i.e., to effectively control the variance attributable to all factors in the test situation except the factor being measured. Such control is usually established by creating uniform testing conditions. In this way potentially confounding factors are either eliminated, or their effect upon the variance is roughly equalized for all subjects; hence, any differences among individuals' test scores are the result of differences in the factor being measured. Thus, feedback on response correctness is thought to be controlled through a uniform system of non-feedback from the tester.

This writer's several years in directing a diagnostic/prescriptive mathematics clinic has led to questioning this procedure in diagnostic test situations, particularly situations involving individually administered tests. During diagnostic testing in the clinic, it has often seemed obvious that some students made conclusions about whether or not their answers were correct, even though no clues were given by the testers. When testers did accidentally give feedback, students' facial expressions of disbelief and comments like "I knew it" further reinforced this observation. If the purpose of withholding correctness feedback is to provide uniform student ignorance of response correctness or a uniform effect from such feedback, then this procedure seems to have failed. In the absence of tester feedback, some, if not all, students seem to substitute their own idiosyncratic systems of feedback.

Support for this observation can be found in both theoretical formulations and limited research findings. Ammons (1956), in reviewing the research literature on knowledge of performance, proposed that there is always some knowledge of performance available to the human performer, ranging from the intrinsic feedback of the senses or an "it feels right" feeling to the extrinsic feedback of evaluative comments like "you are doing well." Stated in other words, Annett (1969), indicated that when extrinsic feedback such as accuracy scores is withheld, most aspects of some intrinsic feedback system remain intact. Smith and Smith (1966), in applying this idea more to the usual learning scene, stated that as an individual's body of knowledge grows, the individual establishes intrinsic standards of accuracy, logic and consistency against which to monitor his responses; thus extrinsic signals like "That's right!" or "Correct!" may serve only to confirm the individual's intrinsic monitoring.

Several research studies have also suggested students' use of subjective feedback. Arps (1917; 1920), in a study of individuals in a finger weight lifting task, found that when experimenter feedback on performance was discontinued, some performers reported imagining pen marks and the physical characteristics of each lift to help maintain their performance. Book and Norvelle (1922) studied children performing simple tasks like writing "a's" as rapidly as possible; when the children were no longer given feedback on their performance, many children were reported to have developed methods of judging how well they did. When Ross (1927) found no difference in performance between college students receiving quiz feedback and students not receiving feedback, he reanalyzed the situation and found the no-feedback students guessed their quiz scores quite well; presumably they received some sort of subjective feedback. In a study of students' confidence in judging whether word pairs were synonyms, antonyms or unrelated, Adams and Adams (1958) found that without training there was only about a 13 percent discrepancy between confidence calls and the actual percent of correct responses, again suggesting some sort of suggestive feedback system. Finally, Ammons (1956) described a study in which subjects were asked to draw lines of a given length, without receiving feedback on their accuracy; subjects began to demonstrate increasing self-consistency around an "incorrect" line length. This suggests the use of subjective feedback relative to some intrinsic standard, even though inconsistent with the experimenter's standard.

The primary reason students are not told whether their diagnostic test responses are correct is to control any effect such feedback might have on test results. It has been argued, however, that when such corrective feedback is withheld, students may rely on subjective feedback. Therefore, instead of providing a uniform effect upon test score variance, it is reasonable that there is a quite uneven effect upon test score variance, depending upon the nature or accuracy of the subjective feedback systems employed and individuals' personalities as they relate to that feedback. If performance on and, therefore, the accuracy of diagnostic tests is influenced by any correctness feedback which the student receives, which is better--a haphazard system of intrinsic types of feedback or a uniform system of extrinsic feedback?

The main purpose of this paper is to review and extend existing literature on feedback as it relates to the testing situation, particularly individually administered diagnostic tests in mathematics. In the paragraphs which follow, the existing literature on feedback is restricted to include only that research which is relevant to diagnostic testing. This research is then presented, along with its implications for educational practice. Finally, areas for further research are identified and a research study currently in progress is described.

Limiting the Review

Massive amounts of research involving feedback have been conducted in the areas of programmed instruction, social reinforcement, test administration, persistence or frustration theory, personality, social psychology, incentives, congruency and dissonance theories. Unfortunately, not all of this research is relevant to the diagnostic testing situation.

The literature on feedback can be divided according to the three types of feedback usually identified: knowledge of response (indication of correct or incorrect), knowledge of correct response (statement of the correct response) and correctional feedback (statement of correct response and reason for its correctness). Since only correct/incorrect feedback is of concern in the diagnostic testing situation, this review was confined to research studies involving knowledge of response feedback. In less current literature this is referred to as knowledge of performance.

A large number of studies have examined the effect of knowledge of response feedback on subsequent task performance. Many of these, however, are not relevant to the diagnostic testing situation. For example, Dweck and Bush (1976) examined the performance of fifth-grade students receiving "failure" feedback. In this study students were told on each of four trials to complete 20 digit-letter substitution problems in one minute; all students were stopped after 15 problems and told they hadn't done well. The results of this study indicated that such feedback led to some or no improvement in performance, but not to a decrement in performance. If findings on other studies were similar to this, one might conclude that telling a child he didn't do very well on a diagnostic test item either has no effect or increases performance, but does not depress it. While this may be an appealing conclusion, one major difficulty with the research exists; the task being performed was not similar to the academic tasks required in the usual diagnostic tests, especially those in mathematics.

Like the study above many knowledge of response feedback studies suffer from this difficulty. As Means and Means (1971) pointed out, much of this research is limited to simple physical, verbal or computational tasks that are not comparable to the complex demands of an academic testing situation. It is difficult to argue, for example, an equitable comparison between the tasks on the Key Math Diagnostic Arithmetic Test (Connolly, Nachtman & Pritchett, 1971) and digit-symbol (Fremont, Means, 1970; Stern, 1972; Dweck & Bush, 1976), auditory discrimination (Dahle & Daly, 1972; Vianello & Evans, 1968), paired-associate (Van de Riet, 1964) or button-pressing (Schmeck & Bruning, 1970) tasks. This is important not only because of the different cognitive demands of those tasks but also because subjects' scores can be affected by whether these tasks are perceived as educationally important (Katz & Greenbaum, 1963).

A similar difficulty in many knowledge of response feedback studies has involved the use of multiple performances on a single task rather than performances on a variety of related tasks. In other words, while many feedback studies involve repeated performance of the same tasks (e.g., digit-symbol substitution) and therefore the same cognitive skill(s), academic testing situations have ordinarily involved performance on a variety of tasks reflecting several cognitive skills. This is important because knowledge of response feedback may play an instructional role when a nearly identical task is repeated directly after feedback. Since academic tests, particularly diagnostic ones, are usually arranged so that knowledge of response feedback on one task should provide little direct help in performing a subsequent task, knowledge of response feedback studies involving multiple performances on a single task are not relevant to the diagnostic testing situation.

In summary, then, the research reviewed in this paper was limited to the knowledge of response literature relevant to the diagnostic testing situation, i.e., studies involving a variety of educationally relevant tasks.

The Literature and Its Implications

A number of research studies involving feedback have been identified which are relevant to the diagnostic testing situation. While some of these have investigated knowledge of response feedback and its effect on test performance, others have examined intra-personal factors like anxiety as they relate to feedback and performance.

The study of knowledge of response feedback as it relates to test performance has been approached in two different ways--by exploring the effect upon test performance of feedback on (a) prior test performance and (b) prior test item performance. In the first approach studies examining the test performance effect of feedback on prior test performance have generally been designed such that students are tested, randomly given sham feedback on their performance, and then tested again. In one such study Bridgeman (1974) examined the effect on a scholastic aptitude test of feedback on a previously administered and obstensibly similar test. For one-third of the students, feedback consisted of a high score and the comment "Excellent! Your problem solving ability is among the best of all seventh-grade students," for another third a low score and the comment "Poor. Your problem solving ability is among the worst of all seventh-grade students," and for the last third no comment. The results indicated that students given positive feedback scored significantly higher than students given negative feedback; no-feedback students scored about the average of the two feedback groups. Gordon and Durea (1948), in a similar study with eighth-grade students, administered two forms of the Stanford-Binet Intelligence Test. Between the tests students received several IQ-type test items. Some of the students received no feedback, while others were told they did poorly both on the between test items and at uniform intervals during the second test. The results showed that the negative-feedback students scored significantly lower than the no-feedback group. Other similar studies have generally found that students who are praised or reproved for their performance subsequently perform about the same or better than students receiving no feedback (Benton, 1936; Blankenship & Humes, 1938; Bornstein, 1968; Gates & Rissland, 1923; Gilchrist, 1916; Hurlock, 1924; Hurlock, 1925; Klugman, 1944; Schmidt, 1941; Tiber & Kennedy, 1964).

Referring to studies like these, Anastasi (1976) indicated that knowledge of response feedback which includes evaluative comments about students' performance or ability has a motivational effect, perhaps operating through the goals which subjects set for themselves in subsequent performance. It appears then that such knowledge of response feedback probably affects students' subsequent test performance; however, the direction in which that performance is affected is not clear. For teachers this suggests that feedback which includes evaluative comments about the students' performance or ability should be withheld before and during test administrations, diagnostic or otherwise. Finally, since knowledge of response feedback may not have the motivational effect suggested above, it is not known whether such feedback in diagnostic testing would influence test performance.

It should be noted that in each of the studies cited above students were randomly given sham knowledge of results feedback on their prior test performance. No doubt many students received feedback which was contrary to their expected performance. What effect this discrepancy between students' expected and reported performance may have had on the results of these studies is not clear; however, one possible explanation is presented later.

Only three studies were identified which examined the effect upon test performance of providing students with immediate item-by-item knowledge of response feedback. Unlike the previous studies, none of these used sham feedback; rather, the feedback given students reflected their actual performance. The earliest of these studies (Morgan & Morgan, 1935) identified three possible effects of providing immediate awareness of response correctness: (a) producing no appreciable modification of performance, (b) increasing effort, attention and critical observation and thereby improving performance, and (c) depressing performance through discouragement. In this study undergraduate psychology students were given a true-false test on "learning." Half of the group used a self-scoring answer sheet and a scoring device which allowed immediate knowledge of correctness; the other half used a mimeographed form of the test with delayed feedback. In another study Angell (1949) used a punchboard device to provide immediate feedback to undergraduate students for each item on mid-semester quizzes in chemistry. Compared to a control group which received delayed feedback, the immediate-feedback group scored significantly higher on the course final examination. Unfortunately, no comparative data was available on the groups' quiz performance, and other intervening factors may have accounted for the difference in examination performance. Finally, Beeson (1971) used a punchboard device to provide immediate item-by-item feedback during mid-semester quizzes and a final examination. The sample consisted of three groups of students--a university class in mathematics for elementary teachers, a university class in remedial mathematics, and a junior high school class in general mathematics. During each of the ten quizzes and one final examination, students randomly received immediate feedback after each item on one-half of the multiple-choice items and delayed feedback on the remainder; no student received immediate feedback on the same half of any two consecutive tests. Care was taken in designing the tests so that feedback on any one item would not cue the answer to an item elsewhere in the test. Except for the final examination, no significant differences were found between test performances under immediate and delayed feedback conditions. There was a slight trend, however, in favor of the immediate feedback test performance.

Based upon these studies, it would appear that immediate item-by-item knowledge of response feedback does not depress (and may even improve) student performance on forced-choice (true-false, multiple-choice) academic tests. Furthermore, there is some indication that, for whatever the reasons, such immediate feedback may increase performance on subsequent tests over the same content. Since, however, diagnostic tests do not usually involve forced-choice items, such conclusions may not be generalizable to the diagnostic testing situation. Receiving knowledge of results feedback from a person, rather than an impersonal testing device, may cast even further doubt upon generalizing these conclusions to individually administered diagnostic testing.[1]

A number of studies have examined various intra-personal factors as they interact with knowledge of results feedback and performance. One such intra personal factor that has been associated with test performance is anxiety. Feldman and Sullivan (1971) and Rosenzweig (1974) have reported that students with lower test anxiety performed better than students with high test anxiety. Cohen (1972) found a relationship between increases in college students' test anxiety and decreases in their verbal/numerical aptitude test performance.

Several studies have investigated the effect of feedback on anxiety. In Cohen's study, he found that feedback indicating performance discrepant with students' expected performance stimulated test anxiety. Feldman and Sullivan (1971) reported that elementary school children's test anxiety was unaffected by positive verbal reinforcement of the

---

[1]Bernstein (1956) found the mere presence of the tester influenced scores on the Thematic Aperception Test.

first correct response on an intelligence test. Finally, McMahon (1973), in a study of the effects of feedback on test anxiety, found that receiving detailed knowledge of test results increased college students' test anxiety.

It would appear from these studies that knowledge of response feedback on test performance probably leads to increased anxiety which, in turn, impedes subsequent performance. However, only Cohen included all three factors--anxiety, knowledge of response feedback and performance--in a single study, and that study only suggests that feedback discrepant with students' expected performance leads to increased anxiety and decreased subsequent performance. Although Cohen's study involved feedback in the form of scores rather than evaluative comments like "excellent" or "poor," this feedback was from a previous test, not item-by-item feedback on the same test. Therefore, the relationship between immediate item-by-item knowledge-of-results feedback, anxiety and test performance remains unclear.

In an earlier discussion it was noted that research procedures involving sham feedback on previous test performance inevitably led to discrepancies between students' expected feedback and that which was reported. Since such discrepancies occurred within groups receiving sham positive feedback as well as negative feedback and since such discrepancies seem to lead to poorer subsequent performance, this may explain the lack of consistent findings in studies of feedback containing evaluative comments.

Further examination of the Cohen study suggests a possible second intra-personal factor--expectation. Although this may only be another way of looking at the effects of anxiety, it appears to have potential. Bridgeman (1974) indicated that self-expectancy deserves more attention from researchers, and Kulhavey (1976), in the context of instruction, stated that how students' expectations influence their use of feedback is a prime area for future research.

No other studies involving intra-personal factors were identified which are relevant to the diagnostic testing situation. However, several studies have examined intra-personal factors as they interact with feedback and performance on non-academic tasks. These are presented to identify factors which may hold promise for future research.

One such intra-personal factor which has been examined in relation to knowledge of results feedback and non-academic test performance is self-concept. Using sham feedback on a test reflecting "sensitivity to other people," Shrauger and Rosenberg (1970) studied the effects of feedback on college students' subsequent performance on a digit-symbol task. Task performance following poor-sensitivity-to-others feedback was generally worse than performance following high-sensitivity-to-others feedback; however, this was found to be mainly attributable to the consistency of this feedback with students' level of self-esteem. Specifically, Shrauger and Rosenberg found that high self-esteem students receiving positive (high-sensitivity) feedback improved their performance, while the reverse was true for low self-esteem students receiving negative (poor-sensitivity) feedback. The performance of high self-esteem students receiving negative feedback and low self-esteem students receiving positive feedback was unchanged. In a similar study, Stern (1972) examined the effect of sham feedback on digit-symbol task performance upon subsequent digit-symbol task performance. Once again task performance following negative feedback was generally worse than performance following positive feedback. Unfortunately, Stern found the reverse relationship between self-concept and feedback; positive feedback more greatly affected low self-concept students, while negative feedback more greatly affected high self-concept students.

As described above, the interaction between knowledge of results feedback, self-concept and performance was examined in only two identified studies and their results conflicted. Therefore, it remains unclear as to whether self-concept is helpful in explaining any relationship between feedback and academic or non-academic test performance.

The results of several studies have suggested other intra-personal factors which may be helpful in explaining any interaction between knowledge of response feedback and non-academic performance. Research in the area of "locus of control" has indicated that when receiving negative feedback from adults, girls tend to attribute failure to a lack of ability (Dweck & Bush, 1976; Dweck & Reppucci, 1973; Nicholls, 1975) and to show a decline in subsequent performance (Dweck & Bush, 1976; Dweck & Gillard, 1975; Maccoby, 1966; Nicholls, 1975; Veroff, 1969); boys, on the contrary, were shown to attribute failure to controllable factors like lack of effort and to display improved subsequent performance. In the area of introversion/extroversion, Fremont, Means and Means (1970) found introverts given negative feedback displayed significantly more anxiety than extroverts, which presumably would lead to a decline in subsequent task performance.

Based upon these studies, one might conclude that knowledge of response feedback has a differential effect on student task performance, depending upon students' locus of control or introversion/ extroversion. The locus of control research might even lead one to conclude that the effect of feedback varies with the students' sex. None of these studies included academic tasks or immediate item-by-item feedback. Therefore, while such factors hold promise for future research, such conclusions may not be generalizable to the diagnostic testing situation.

Future Research

The existing research on knowledge of response feedback as it relates to the testing situation, as presented in this paper, has led to three conclusions:

(b) immediate item-by-item knowledge of response feedback does not depress (and may improve) student performance on forced-choice academic tests; and

(c) knowledge of response feedback on previous test performance which is discrepant with students' expected performance probably leads to increased anxiety and decreased subsequent performance.

Conclusions with respect to diagnostic testing were lacking because of the general lack of correspondence between the research conditions and the diagnostic testing situation. Aside from performance on non-academic tasks, factors leading to this lack of correspondence included knowledge of response feedback which included evaluative comments like "excellent" or "poor," forced-choice academic test formats like true-false or multiple-choice, and feedback provided by a device rather than a person.

In light of the previous research (or lack of it), there are three areas of needed research. Stated in the form of questions, these are:

(1) <u>In the diagnostic testing situation, does immediate item-by-item knowledge of response feedback affect performance on a diagnostic test?</u>

Previous research has suggested that such feedback has little effect upon performance on forced-choice academic tests. To answer this question, further research needs to be conducted in which test items are free-response and feedback is provided by the tester, rather than by a mechanical device.

(2) <u>In the diagnostic testing situation, which intra-personal factors, if any, help explain the relationship between immediate item-by-item knowledge of response feedback and diagnostic test performance?</u>

Although rather inconclusive, previous research has suggested that the relationship between feedback and subsequent performance may not be a simple one, but is influenced by personal factors. Anxiety appears to be a highly probable factor to help explain this relationship. Other potential factors suggested in the literature include student expectations, self-concept, locus of control and introversion/extroversion.

(3) <u>In the diagnostic testing situation, does students' intrinsic, subjective knowledge of response feedback affect their performance on a diagnostic test?</u>

It was proposed earlier in this paper and supported by some research that, in the absence of tester feedback, students substitute their own idiosyncratic systems of feedback. However, with the possible exception of Ross (1927), no research was identified which dealt with subjective feedback as it relates to performance on academic or diagnostic tests. Because of the total void of research in this area, a series of sub-questions needs to be answered: When the tester withholds feedback during a diagnostic test, do students provide themselves with knowledge of response feedback? How accurate is this feedback? Does it affect diagnostic test performance? Does this feedback interact with tester-provided feedback to affect performance on a diagnostic test? The small amount of research on the discrepancy between feedback and expected performance relates to this last sub-question, suggesting the likely possibility of such an interaction.

A Study

Based upon a review of the literature on feedback as it relates to diagnostic testing, three major questions or areas of research were posed. To help provide answers to these questions, a research study was recently conducted by the author.

The purposes of the study were (1) to determine if item-by-item knowledge of response feedback (induced self, tester-provided, both) affected students' performance on a diagnostic mathematics test, (2) to explore the accuracy of students' intrinsic feedback as reflected in their self-reported knowledge of response feedback on a diagnostic mathematics test, and (3) to examine the relationship between students' performance expectations and performance on a diagnostic mathematics test. Organized by their appropriate general questions, the null hypotheses of the study are formulated below:

(1) Does item-by-item knowledge of response feedback affect performance on a diagnostic mathematics test?

- (a) No significant differences in diagnostic mathematics test scores will be observed between students receiving no knowledge of response feedback and students receiving item-by-item (induced self, tester-provided, induced self followed by tester-provided) knowledge of response feedback.

(2) In the diagnostic test situation, do student expectations influence the effect item-by-item knowledge of response feedback has upon diagnostic test performance?

- (a) Differences in the diagnostic mathematics test performance of students expecting high and low test performance will not differ significantly between students receiving item-by-item knowledge of response feedback.

(3) How accurate is students' induced self feedback on a diagnostic mathematics test?

- (a) Students' actual scores on a diagnostic mathematics test will not differ significantly from students' scores as indicated by their item-by-item induced self feedback.

Method

Subjects

The sample consisted of 98 fourth-grade and 94 sixth-grade students from the Washington Elementary School District, Phoenix, Arizona. These students represented four classrooms at each grade level randomly selected from district classrooms. None of the classrooms theoretically represented a homogeneous ability group in math, reading or any other area.

Materials

Two specially-constructed diagnostic mathematics tests were used in this study, one for each grade level. These tests were designed to survey the major mathematics content for the grade level in which it was used. Both tests were divided into two twenty-item parts, with corresponding items in each part testing the same mathematics concept or skill. Test items in each part were arranged in ascending fashion, with earlier items testing concepts or skills which (a) appeared logically prerequisite to those tested in later items, or (b) are normally taught prior to those tested in later items. Within each twenty-item part, care was taken so that feedback on any one item would not cue the answer to a subsequent item. All test items were free response, and each item was printed on a separate five-by-eight inch page. Thus, at each grade level, the diagnostic mathematics test consisted of two parallel sub-tests, each surveying the same major mathematics concepts and skills for that grade level and each arranged in ascending order.

The two part test was used because, whatever effect feedback might have in the first part of the test, it was felt that this effect would be even more apparent in the second part. For example, if feedback produces performance-debilitating anxiety, then not only should test performance on the first part be depressed (compared to a non-feedback group), but test performance on the second part should be even further depressed (compared to a non-feedback group).

Procedure

Students from each classroom in the sample were randomly selected in groups of five, and each of these groups were in turn randomly assigned to one of four feedback treatment groups--one to receive no feedback, one to receive knowledge of performance feedback from the tester, one to indicate after completing each item whether they expected their response to be correct or incorrect (induced self feedback), and one to receive induced self feedback followed by tester feedback. An attempt was made to assign students so that all four classrooms at each grade level were equally represented in each treatment group. All students were tested by the same tester in a room other than the usual classroom.

As the testing began, the students were told that they were going to take a two-part test which would help identify any areas of mathematics in which they might need some help. Since no data had been previously collected on the equivalence of the two test parts, half of the students in each treatment group were given one part first and the other half the second part first. After writing their name and teacher on the front cover, the students were instructed to briefly look through the 20 items and write the number of items they expected to answer correctly.

When this was done, students received different instructions, depending upon to which of the four feedback groups they had been randomly assigned. In the first group (the non-feedback group), students were told to answer each item and then mark an "X" in the box to the right of the problem to show it had been finished. Students in the second group (the induced self feedback group), were instructed to answer each item and then mark one of the following words: "RIGHT," "right," "wrong" or "WRONG." They were to mark one of the capitalized words if they felt certain their response was right or wrong, and mark one of the words in lower-case if they felt their response was probably right or wrong. After responding to each test item, students in the third group (the tester-provided feedback group) were told to mark an "X" in the box to the right of the problem to show it had been finished. The tester than pointed to the word "RIGHT" or "WRONG" next to the problem, indicating the correctness of the students' responses. Finally, in the fourth group (the induced self and tester-provided feedback group), students were told to answer each item and mark one of the words "RIGHT," "right," "wrong" or "WRONG" to show how certain they were of their response correctness. The tester then pointed to the word "RIGHT" or "WRONG" to indicate whether or not the students' response was correct. The students in all feedback groups were not allowed to proceed from one item to the next until all five members of the testing group had completed the item.

When the students had completed the first part of the test, they were given a one or two minute break and then proceeded on to the second part. As with the first part, they were instructed to briefly look through the items and write the number of items they expected to answer correctly. The students then completed the test, following the same instructions as in the first part.

Thus, in randomly selected groups of five, fourth and sixth-grade students were administered a two-part, ascending diagnostic mathematics test. Before beginning each part of the test, students indicated the number of items they expected to answer correctly. As they completed each test item, students either received no feedback, self feedback, tester-provided feedback, or both self and tester-provided feedback. Over a four week period, two weeks for each grade level, each test administration took approximately 40 minutes.

Table 1

Descriptive Statistics for Diagnostic Mathematics Test Performance

Feedback Group	Grade	n	Part A				Part B			
			Expected		Actual		Expected		Actual	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
No Feedback	4	29	17.34	2.52	13.76	3.53	16.93	2.31	14.14	3.47
	6	22	16.41	2.32	11.90	4.93	15.54	3.79	11.68	4.60
Induced Self Feedback	4	25	17.00	3.00	13.48	2.94	17.12	3.00	14.32	2.82
	6	26	15.42	3.32	10.62	3.97	15.54	3.56	10.73	3.92
Tester Feedback	4	24	17.17	3.28	12.54	3.96	15.58	3.05	13.62	3.47
	6	23	16.04	3.76	10.70	4.10	13.48	3.91	10.96	4.00
Induced Self-Tester Feedback	4	20	17.40	4.22	13.00	3.92	15.60	3.12	13.90	3.29
	6	23	16.13	3.22	10.30	3.88	12.39	3.82	10.48	4.39

Table 2

Two-Way Analysis of Variance of Diagnostic Test Performance

Grade	Test Part	Source of Variation	SS	df	MS	F
4	A	Main Effects	123.14	4	30.79	2.58*
		Feedback Group	35.52	3	11.84	.99
		High/Low Expectancy	101.05	1	101.05	8.46**
		2-Way Interactions				
		Group x Expectancy	33.41	3	11.14	.93
		Explained	156.53	7	22.36	1.87
		Residual	1075.05	90	11.94	
		Total	1231.60	97	12.70	
	B	Main Effects	65.52	4	16.38	1.59
		Feedback Group	13.22	3	4.41	.43
		High/Low Expectancy	58.85	1	58.85	5.72*
		2-Way Interactions				
		Group x Expectancy	27.71	3	9.24	.90
		Explained	93.23	7	13.32	1.30
		Residual	925.76	90	10.29	
		Total	1018.99	97	10.50	
6	A	Main Effects	136.15	4	34.04	2.23
		Feedback Group	40.66	3	13.55	.89
		High/Low Expectancy	102.66	1	102.66	6.74*
		2-Way Interactions				
		Group x Expectancy	172.70	3	57.57	3.78*
		Explained	308.86	7	44.12	2.90**
		Residual	1310.35	86	15.24	
		Total	1619.20	93	17.41	
	B	Main Effects	168.95	4	42.24	2.66*
		Feedback Group	43.93	3	14.64	.92
		High/Low Expectancy	132.98	1	132.98	8.36**
		2-Way Interactions				
		Group x Expectancy	105.42	3	35.14	2.21
		Explained	274.37	7	39.20	2.46
		Residual	1368.18	86	15.91	
		Total	1642.55	93	17.66	

*$p < .05$
**$p < .01$

Table 3

Comparison of Students' Actual Diagnostic Test Performance and Performance Reflected in Students' Induced Self Feedback

Feedback Group	Grade	n	Score	Part A Mean	Part A SD	df	t	Part B Mean	Part B SD	df	t
Induced Self			Self	17.56	1.94			17.52	1.74		
Feedback	4	25	Actual	13.48	2.95	24	-7.45**	14.32	2.82	24	-5.72**
			Self	16.00	2.74			16.12	3.20		
	6	26	Actual	10.62	3.97	25	-6.91**	10.73	3.92	25	-7.44**
Induced Self-			Self	16.70	2.66			16.55	2.48		
Tester Feedback	4	20	Actual	13.00	3.92	19	-4.65**	13.90	3.29	19	-4.36**
			Self	14.78	4.17			13.39	4.81		
	6	23	Actual	10.30	3.88	22	-6.80**	10.48	4.39	22	-3.69**

**$p < .01$

Results

Measures of equivalent-forms test reliability were obtained for the two forms of the fourth-grade diagnostic mathematics test and the sixth-grade test. To avoid the possible effects of feedback upon the reliabilities, only scores from students receiving no feedback were used; half the students at each grade level had received one form first, and half the other form first. The coefficients of equivalence for the fourth and sixth-grade tests were respectively .8475 and .9520.

Students' expected and actual performance on the diagnostic mathematics tests are displayed in Table 1. A 4 x 2 factorial design was used to ascertain the effects of item-by-item knowledge of response feedback for each part of the diagnostic mathematics tests. At both grade levels multivariable analysis of variance indicated no significant differences in performance between students receiving no feedback on the diagnostic mathematics test and students receiving knowledge of response feedback ($\underline{F}$=1.15, $\underline{df}$=2/93 and $\underline{F}$=1.54, $\underline{df}$=2/89 for fourth and sixth-grades respectively). Because of the overall lack of significance, univariate $\underline{F}$'s were not computed.

A two-way analysis of variance procedure was used to ascertain whether student's performance expectations influenced the effect of knowledge of response feedback on student's diagnostic mathematics test performance. Students were separated into high and low performance expectancy groups according to the mean expected performance on the first part (Part A) of the diagnostic mathematics tests at each grade level (17.22 and 15.98 for grades four and six respectively). The actual test performance of students expecting high and low performance was then compared for the four feedback groups. The results of the analyses of variance are presented in Table 2. The only significant interaction was in Part A for sixth-grade students. Since one would expect this significance to be maintained for Part B, it may have been a chance occurrence.

The accuracy of students' scores as indicated by their item-by-item induced self feedback was checked by comparing these scores to students' actual performance. A summary of these scores and analyses is shown in Table 3. For both grade levels and for both feedback groups involving induced self feedback, students consistently and significantly self-reported greater performance than that actually demonstrated. It is interesting to note that the difference remained significant in the second part of the diagnostic test, even though two groups of students generally received some disconfirmation of their self-feedback from the tester.

Discussion

The present study explored questions relative to item-by-item knowledge of response feedback as it relates to diagnostic testing in mathematics. The results would seem to lend support to the argument raised in the introduction that item-by-item knowledge of response feedback does not affect performance on a diagnostic mathematics test. This conclusion, however, must be qualified.

It was argued that in the absence of tester feedback, students substitute their own subjective feedback. If this is true, then, students designated as receiving no feedback in this study were actually receiving subjective feedback. Thus, this conclusion must be qualified to indicate that item-by-item feedback affects students' diagnostic test performance no differently than students' subjective feedback.

The results of this study seem to further suggest that students' performance expectations do not influence the effect of knowledge of response feedback on diagnostic test performance. A significant interaction between students' expectancy and feedback was found for sixth-grade students on Part A of the diagnostic test, but this appears to be likely the result of chance. Performance expectancy, therefore, does not appear to be an intra-personal factor which interacts with knowledge of response feedback in affecting students' diagnostic mathematics test performance.

In the present study students' scores on the diagnostic mathematics tests, as reflected in their self-reported item-by-item knowledge of response feedback, differed significantly from their actual scores. Students consistently self-reported performance which was greater than their actual performance. This suggests that students' induced self feedback on a diagnostic mathematics test is relatively inaccurate. If students' subjective knowledge of response feedback is reflected in their self-reported feedback, as induced in this study, then students' subjective feedback is inaccurate. The fact that students tended to rate themselves as doing better than they actually were supports Cronbach's (1960) observation that children are often enormously pleased with very inferior responses.

The prior review of the literature on knowledge of response feedback revealed no conclusions with respect to the diagnostic testing situation. One purpose for conducting the present study was to begin empirically investigating and answering questions concerning the use of such feedback in diagnostic mathematics testing. While the results appear to suggest that providing item-by-item knowledge of response feedback on diagnostic mathematics tests affects students' performance no more than withholding such feedback, intra-personal factors other than performance expectancy may exist (e.g., anxiety or self-concept) which indicate differential effects of feedback on diagnostic test performance. Questions concerning students' intrinsic, subjective knowledge of response feedback in the diagnostic testing situation remain largely unanswered; this study has only suggested that such feedback may be relatively inaccurate.

In this paper the issue of providing knowledge of response feedback in diagnostic mathematics testing has been raised. Although an initial attempt was made to investigate several questions relevant to knowledge of response feedback in diagnostic testing, a number of questions remain, and further research needs to be conducted before the issue will be resolved.

REFERENCES

Ammons, R. B. Effects of knowledge of performance: A survey and tentative theoretical formulation. Journal of General Psychology, 1956, 54, 279-299.

Anastasi, A. Psychological testing (4th ed.). New York: Macmillan, 1976.

Angell, G. W. The effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. Journal of Educational Research, 1949, 42, 391-394.

Annett, J. Feedback and human behavior. Baltimore, Maryland: Penguin, 1969.

Arps, G. F. A preliminary report on work with knowledge versus work without knowledge of results. Psychological Review, 1917, 24, 449-455.

Arps, G. F. Work with knowledge of results versus work without knowledge of results. Psychological Monographs, 1920, 28, No. 125.

Beeson, R. O. Immediate knowledge of results and test performance. Paper presented at the meeting of the American Educational Research Association, New York, February 1971. (ERIC Document Reproduction Service No. ED 048 375).

Benton, A. L. Influence of certain incentives upon intelligence test scores of school children. Journal of Genetic Psychology, 1936, 49, 494-496.

Bernstein, L. The examiner as an inhibiting factor in clinical testing. Journal of Consulting Psychology, 1956, 20, 287-290.

Blankenship, M. V. & Humes, J. F. The effect of praise and reproof upon memory span performance. American Journal of Psychology, 1938, 51, 527-531.

Book, W. F. & Norvelle, L. The will to learn: An experimental study of incentives in learning. Pedagogical Seminary and Journal of Genetic Psychology, 1922, 29, 305-362.

Bornstein, A. V. The effects of examiner approval and disapproval upon the performance of subjects on the performance scale of the Wechsler Intelligence Scale for Children. (Doctoral dissertation, Florida State University, 1967.) Dissertation Abstracts International, 1968, 28, 3047B. (University Microfilms No. 68-339)

Bridgeman, B. Effects of test score feedback on immediately subsequent test performance. Journal of Educational Psychology, 1974, 66, 62-66.

Cohen, R. M. Effects of feedback on test anxiety and performance as a function of certain personal characteristics. (Doctoral dissertation, New York University, 1971.) Dissertation Abstracts International, 1972, 32, 6094B. (University Microfilms No. 72-13, 342)

Connolly, A. J., Nachtman, W. & Pritchett, E. M. Key math diagnostic arithmetic test. Circle Pines, Minn.: American Guidance Service, Inc., 1971.

Cronbach, L. J. Essentials of psychological testing (2nd ed.). New York: Harper & Brothers, 1960.

Dahle, A. J. & Daly, D. A. Influence of verbal feedback on auditory discrimination performance of mentally retarded children. American Journal of Mental Deficiency, 1972, 76, 586-590.

Dweck, C. S. & Bush, E. S. Sex differences in learned helplessness: I. Differential debilitation with peer and adult evaluators. Developmental Psychology, 1976, 12, 147-156.

Dweck, C. S. & Gillard, D. Expectancy statements as determinants of reactions to failure: Sex differences in persistence and expectancy change. Journal of Personality and Social Psychology, 1975, 32, 1077-1088.

Dweck, C. S. & Reppucci, N. D. Learned helplessness and reinforcement responsibility in children. Journal of Personality and Social Psychology, 1973, 25, 109-116.

Feldman, S. E. & Sullivan, D. S. Factors mediating the effects of enhanced rapport on children's performance. Journal of Consulting and Clinical Psychology, 1971, 36, 302.

Freemont, T., Means, G. H. & Means, R. S. Anxiety as a function of task performance feedback and extraversion-introversion. Psychological Reports, 1970, 27, 455-458.

Gates, G. S. and Rissland, L. Q. The effect of encouragement and discouragement upon performance. Journal of Educational Psychology, 1923, 14, 21-27.

Gilchrist, E. P. The extent to which praise and reproof affect a pupil's work. School and Society, 1916, 4, 872-874.

Gordon, L. V. & Durea, M. A. The effect of discouragement on the revised Stanford-Binet Scale. Journal of Genetic Psychology, 1948, 73, 201-207.

Hurlock, E. B. The value of praise and reproof as incentives for children. Archives of Psychology, 1924, 11 (71), 78.

Hurlock, E. B. An evaluation of certain incentives used in schoolwork. Journal of Educational Psychology, 1925, 16, 145-149.

Katz, I. & Greenbaum, C. Effects of anxiety, threat, and racial environment on task performance of Negro college students. Journal of Abnormal and Social Psychology, 1963, 66, 562-567.

Klugman, S. F. The effect of money incentives versus praise upon the reliability and obtained scores of the revised Stanford-Binet test. Journal of General Psychology, 1944, 30, 255-269.

Kulhavey, R. W. Feedback in written instruction. (Tech. Rep. 1) Tempe, Arizona: Arizona State University, Laboratory for the Study of Human Intellectual Processing, February 1976.

Maccoby, E. E. Sex differences ' intellectual functioning. In E. E. Maccoby (Ed.), The development of s ifferences. Stanford, California: Stanford University Press, 1966.

McMahon, M. P. Effects of knowledge of ability test results on academic performance and test anxiety. Journal of Counciling Psychology, 1973, 20, 247-249.

Means, R. S. & Means, G. H. Achievement as a function of the presence of prior information concerning aptitude. Journal of Educational Psychology, 1971, 62, 185-187.

Morgan, C. L. & Morgan, L. V. Effects of immediate awareness of success and failure upon objective examination scores. Journal of Experimental Education, 1935, 4, 63-66.

Nicholls, J. G. Causal attributions and other achievement-related cognitions: Effects of task outcomes, attainment value, and sex. Journal of Personality and Social Psychology, 1975, 31, 379-389.

Rosenzweig, S. J. The effects of examiner anxiety level, student test anxiety level, and exar ner-student sex interaction, on student performance in a group test-taking situation. (Doctoral dissertation.) DAI, 1974, 35, 1511A.

Ross, C. C. An experiment in motivation. Journal of Educational Psychology, 1927, 18, 337-346.

Schmeck, R. R. & Bruning, J. L. Frustration theory and quality of performance: Elicitation and elimination of competing responses. Psychological Reports, 1970, 26, 987-994.

Schmidt, " O. The effects of praise and blame as incentives to learning. Psy 'ogical Monographs, 1941, 53 (3), 56.

Shrauger S. & Rosenberg, S. E. Self-esteem and the effects of success an ilure feedback on performance. Journal of Personality, 1970, 38, 404- 7.

Smith, K. V. & Smit M. F. Cybernetic principles of learning and educational design. New York: Holt, Rineha t & Winston, 1966.

Stern, D. Effects of level of self-concept, type of feedback, and internal-external locus of control of reinforcement upon performance on a digit symbol task. (Doctoral dissertation, University of Montana, 1972). Dissertation Abstracts International, 1972, 33, 2359B. (University Microfilms No. 72-29, 347)

Tiber, N. & Kennedy, W. A. The effects of incentives on the intelligence test performance of different social groups. Journal of Consulting Psychology, 1964, 28, 187.

Van de Riet, H. Effects of praise and reproof on paired-associate learning in educationally retarded children. Journal of Educational Psychology, 1964, 55, 139-143.

Veroff, J. Social comparison and the development of achievement motivation. In C. P. Smith (Ed.), Achievement-related motives in children. New York: Russell Sage, 1969.

Vitanello, M. A. & Evans, S. H. Note on pitch discrimination learning. Perceptual and Motor Skills, 1968, 26, 576.

Reaction Paper
to
Feedback in Diagnostic Testing

Lloyd Hutchins
Francis Marion College

I would like to begin with a few general reactions and then deal with a particular interpretation of the results which, in my opinion, are important in the planning of the next level of research in this area.

Drs. Romberg and Uprichard, in their paper, discussed classes of research such as discovery-oriented research and research for confirmation. At the end of the continuum we have to have considerable inferential rigor because our purposes are different for discovery. While it is generally an influential mold, it's much more important to consider every possibility and make accommodations to logistical needs that otherwise could not be done. We should not be deterred from pursuing our interests simply because we can't always please the people who spend alot of time shocking rats. In a discovery study, an exploratory investigation, there are some informalities in the analysis but they are perfect y legitimate in this context.

Another point I would like to make has to do with the partitioning of students into what could be considered to be emotional categories for further research, e.g., high-anxiety people, secure people, etc. I think that this is extremely useful and that it will allow us to move more quickly in the next stage.

I would like to address an aspect of research in the second stage that I feel is probably of particular importance. Because it's important in the second stage does not mean it is not important in the stage represented by Dr. Engelhardt's study. That is, I think that his study had to be done in this way before we could do what I'm going to suggest. It's very definitely not a criticism of his procedure, which I think is correct. With this in mind, I would like to pose the following for consideration.

At the next level of research I think we must partition our measurements. We must measure the effects of positive and negative feedback separately. If we measure feedback as a Gestalt, a complete entity, we will not be testing whether any effect exists, but whether the effects cancel each other. Assume we have a situation in which a child gives approximately 50% correct and 50% incorrect responses. That is, a child receives a positive reinforcement in one case and a negative reinforcement in the other. It's probably reasonable to assume that the effects will cancel. If we look, for example, at the table of means and standard deviations, we do have approximately 50% correct response in every instance. Consequently, in the next step we might want to measure the positive and negative reinforcement differently. That is, keep track of the amount of positive treatment and the amount of negative treatment applied.

Another adjustment I feel is important is the control of the reinforcement variable. Rather than taking the average of positive and negative effects in some degree we should compare intervals where we know exactly what the degree of treatment was. This, of course, would require a pre-testing and a fairly elaborate sorting of response on some equivalent pre-test form. But, I think if we intend to make use of the procedure in clinical situations we really need to know the response to a fixed treatment, not the mean effect of some group of people.

Regarding the conclusion, say we did a second study in which the child was being told he/she was wrong about half the time. In addition, let's say there was no difference between the feedback by the tester and the child's own induced feedback. We cannot assume from this that this relationship will hold over the entire range. This may be, and very likely is, a catastrophic phenomenon. When you reach a certain level of stimulus, things happen very quickly. I think that a child who was told he/she was wrong 85% of the time by an authority figure might very well experience some considerable effect different from the child's own induced by anxieties over suspicion he/she was wrong much of the time. Similarly, I think there would be comparable difference in positive reinforcement.

Now, by catastrophic phenomena again, we mean the effects are not linear. We actually need to have fixed treatments over intervals of response of 10% or so, but this would be difficult to do. I think that in the mid-range there isn't going to be a very large difference but that when we move away from the mid-range where, for example, a child is receiving a high percentage of positive reinforcement (or negative reinforcement), there will be some important effect on performance. The point is, we have to have fixed effects, not average effects, and we have to compare tester feedback with self-induced feedback at different intervals before we really have something that will allow us to make clinical decisions.

If the point of providing feedback is to improve our measurement of the cognitive construct - to reduce its being confounded by anxiety or something of that sort, then it is extremely important to use fixed effect and to measure in intervals.

Counting Performance and Achievement
Some Preliminary Observations

Leroy G. Callahan
SUNY/Buffalo

> Obviously, the first skill of any consequence that children must master is the ability to count. ...No system can afford to fail in developing this skill at a rather early point in the child's learning experience. (Buchanan, p. 33)

> The most basic of all skills is that of competence in the use of number systems for counting, comparing, ordering and measuring. (Hilton, p. 88)

> Basic Skills (include)
> A firm understanding of the basic meaning and uses of numbers for counting, comparing and ordering. (Rubenstein, p. 177)

Renewed concerns with basic skills in mathematics have focused some attention on the counting process, as suggested by the quotes above taken from <u>Conference on Basic Mathematical Skills and Learning</u>, Vol. 1 (1975). Associated with these concerns is the interest in early identification of children who may have difficulty in learning basic mathematical skills. The Project for the Mathematical Development of Children (PMDC) group has been involved with examining and describing some aspects of the enumeration process in young children. The psychological processes, especially the role of perception, involved with enumeration have been of interest to psychologists for over a century (example: Jevons, 1871).

This paper explores some of the cognitive processes that appear to be involved with counting, and also reports on observations of first graders involved with performing rote and rational counting tasks. I, and some of my students, have been involved with interviewing small numbers of first graders for the past four years. This year as part of a more comprehensive project we were able to carry out about one thousand interviews of first graders. The data are presently being processed. Observations reported in this paper are based on a smaller number of interviews carried out in previous years. Finally, some speculations will be made on the diagnostic value of these tasks in predicting learning strengths and weaknesses of students within a broader context.

The subset of tasks on the interview concerned with counting involved two aspects, rote counting and rational counting. The former involved the examiner (E) asking the child, "How far can you count?" and then "Can you show me?" The child was also asked to count on from a point in the system other than one, and to count backward from different points in the system. The rational counting tasks included counting 3, 7, 11 and 16 counting blocks from a pile; choosing a dot card with N dots when presented with a trio of dot cards (N, N-1, N+1); a sorting task with dot cards with different arrangements of 5, 7 and 10 dots which were to be sorted onto criterion cards with displays of 5, 7 and 10 dots; and a task requiring ordering of dot cards and numeral cards.

Rote counting can be described as a verbal chain (Gagne, 1970). A series of words are linked in an arbitrary sequence. Through modeling and/or repetitive practice children may become proficient in performing this verbal chain. In the interview observations could be made on the length of the verbal chain, points in the chain that appear to be more difficult, and other factors that may affect performance.

One general observation was that first graders differ in their rote counting performance on entering first grade. Danmark (1975) reported that when (E) terminated the chain at "thirty-five" over 50% of approximately 200 first graders reached that point; about 16% reached "twenty-nine"; about 12% reached "nineteen"; about 9% reached "ten"; and about 10% could not get to "ten" in the chain. In our procedure (E) did not terminate the count until one hundred.

Of about 40 observations made from this batch of students, one could only count to 10 or less; two stopped between 11 and 20; eight between 21 and 30; eight between 31 and 40; one between 41 and 50; three between 51 and 60; and all the rest (18) made it to one hundred.

Although the sample of performance being reported on was small and non-random, it was interesting to note that if a student made it to "sixty" in the chain they then could make it to one hundred. Another observation was that most of the terminal points in the chain were at the point of movement into a new decade, i.e., 29, 39, 49. Further, some (E) [illegible] served that even for those able to make the count to one hundred there was hesitancy and pauses at some of these points. Using analogy from the physical sciences, it appeared that a relatively small amount of ergs of cognitive energy are required to continue the flow of the verbal chain within a decade, but additional ergs are required to cross the threshold to the next decade. In speculating on why students who make it through the fifties could make it all the way to one hundred, one possible explanation might be the linguistic similarity between the decade names sixty, seventy, eighty, ninety and the familiar six, seven, eight and nine. With the exception of forty, there is a degree of dissimilarity between twenty, thirty, forty, fifty (where students were terminating the chain) and the familiar two, three, four and five.

A few further observations on rote counting. There often was a discrepency between how far the children say they can count and how far they can, in fact, count. Typically, the point that they say they can count to is less than where they can count in their performance.

Another rather common observation was the pride so many youngsters took in "showing off" their counting ability. Counting to one hundred may not be a particularly fun thing for older children, but many of these children seemed to enjoy being the soloist in a counting performance.

Although it was an unsystematic inquiry, many of those able to count to one hundred were asked, "Where did you learn to count so well?" The most common responses were associated with older sisters or brothers, parents (usually the mother) or in a few cases grandmothers. Very few indicated the school. Ogletree (1970) alluded to this influence when he stated, "...the intellectual part of rote counting lies in the fact that number words are strictly ordered, and whenever the child muddles up the sequence, or skips a word, his older playmates or adults correct him." Ausubel (1964) commented, a number of years ago, on the importance of furnishing acceptable models of speech and supplying corrective feedback in relation to early learning, especially for the disadvantaged child.

Students were asked to count-on from 6. Our observations were very similar to those reported by Denmark (1975). The majority of students could perform the task. Some had to have the direction clarified by cueing. Relatively few were unable to perform after the directions were clear.

Students were asked to count backward from 10; a number less than ten (6); and a number greater than ten (12 or 13). Of about 40 observations made, about 14 were unable to count from 6, 10, or 12; two were able to count from six but not 10 or 12 [illegible] were able to count from 6 and 10 but not 12; and eighteen were [illegible] nt from 6, 10 and 12. Denmark reported that about 80% of th[illegible] ders could count backward from six once the directions wer[illegible] to them (some needed cueing). We also would cue the studen[illegible] ...like 'ten,' 'nine,' 'eight.'"), yet this group of first-graders reflected somewhat more difficulty than the Florida and Georgia children. Counting backward from twelve was very difficult.

Observations of the backward counting task included the evident lack of experience by many children in any backward counting. Even after cueing by (E) there were many blank expressions. Some students who indicated that they had never counted backwards understood the nature of the task through the cueing and would attempt the task. Some would succeed--usually much more slowly than with forward chains. For some the count backward would begin but at some point a number word would trigger a switch to a forward chain, i.e., "ten, nine, eight, seven, eight, nine...." For some the backward counting chain was quite a complex cognitive process. The chain would begin "ten, nine, eight, (long pause) seven, six, (long pause) five, (long pause) four," etc. The long pause [illegible] ually involved starting at "one" and counting up to the point where the block in the backward sequence occurred, (this could be vocal or subvocal counting) then continuing the backward chain. For some few these "pause processes" occurred for every number name in the backward chain of counting. As suggested, this was quite a complex cognitive for.. For students proficient in the task the backward chain of names seemed to be no more demanding than the forward chain. Generally, our observation has been that significant numbers of entering first-grade children have limited, or no, experience and proficiency in counting backward.

Proceeding to rational counting tasks, the first involved placing a pile of blocks before the child and asking him to show four, eight, ten, and thirteen. The enumeration process involved in such a task has been examined by various investigators. Beckwith and Restle (1966) describe the act of simple enumeration to include (1) the sequenced chant of number names, (2) an associated indicator response (such as pointing) which is synchronized with each number name, (3) a discrimination between the set already counted and the set as yet uncounted, and (4) the termination of the process must be recognized--usually by the emptiness of the uncounted set. In regard to this last point, enumerating a subset of five (for example) from a larger set appears to be more difficult than the exhaustive enumeration of a set of five. Wang, Resnick, and Boozer (1971) hypothesized that the addition of a "memory component" is very likely the factor that increased the difficulty of such a task. Their data suggested that children continued to count out objects beyond the number specified in the instruction. The tasks in the present interview involved the more difficult enumeration of a specified subset from a larger set.

As might be expected, it was observed that demonstrating the two larger numbers, ten and thirteen, was more difficult than demonstrating four or eight. The "memory component" seemed to contribute to the observed difficulties, especially with the larger numbers. The students didn't necessarily go beyond the number requested (although some did), but would often have to stop the enumeration process to ask (E) what number had been requested. With some this might occur more than once during the processing of one of the larger numbers. Some students would pull blocks from the pile, organize them (linear rows or arrays), and then enumerate--putting extras back or pulling more from the pile as needed. Some would pull blocks from the pile with a finger in synchronization with the oral count (vocal or subvocal) with considerable physical separation between the counted subset and uncounted set; others hardly separated the counted blocks from the uncounted set making only a slight physical contact. In the former case it was relatively simple to run a check on the accuracy of the count if the students wished (and many did want to check their count), in the latter case such checks were not possible.

Another interesting observation that could be made involved the way students related the sequencial tasks. After demonstrating a set of four the next instruction was to show eight blocks. Some would quickly relate that request to the previous task and build-on from four. Others did this, but would go back to count the four first, as if to see if it was still four; after the four more were in place some would then go back again and count the entire eight. There seemed to be some distrust of the stability of the number property of the set so that, for safety's sake, they went back and checked quite often. Of course, with some, this again may have involved a "memory component." The synchronization of the number words with the motor indicator response of touching or moving the blocks was another source of error in responses. Again, this generally showed up at the higher number levels. Some of the synchronization problems were gross and were obvious even while enumerating small numbers of blocks; a few were very fine with the timing just slightly off so that the cumulative dissonance between count and indicator movement would not throw the count off until it reached eight or nine.

The next two rational counting tasks involved the use of dot cards in discriminating and recognizing number properties. In each case circular blue stickers were fixed to white 3 x 5 cards to become the stimulus objects. Because it appeared that tasks dealing with small numbers of dots on cards posed different cognitive processing demands than cards with larger numbers of dots, a short digression to examine that line of research is in order.

Psychologists have been intrigued for years with the question dealing with span of apprehension. As early as 1871, Jevons (1871) found that his estimations (of beans cast into a box) were invariably correct when the number of beans was less than 5; as the number of beans increased the percentage of correct estimations decreased--and in fairly regular fashion. Warren (1899) concluded that, except under special stress of attention, or with subjects especially apt in this direction, the function of perception counting (number displays that appear to be apprehended as a whole, without enumeration) is limited to the numbers one, two, and three. To apprehend numbers greater than 4, then, some other perceptual function must come into play. Taves (1941) concluded from his data that there are two mechanisms for the perception of visual numerousness. One operates when the number of dots in the stimulus field is small (numbers up to 6), the second mechanism operates when the number of dots is so great as to prohibit direct and rapid recognition of number (greater than 6). Kaufman et.al. (1949) supported the two different mechanism hypotheses and that the change occurred about at 6 stimulus dots. Their data did not support Saltzman and Garner (1948), and other earlier experimenters, that suggested there was an immediate cognition of number with presentations of 1 to 6 stimulus dots. They found the time between stimulation and report increased regularly with dot presentations from 1 to 6, therefore bringing into some question the "immediate cognition" hypotheses. A considerable number of independent variables have been examined as to their influences on the apprehension question. Some of these variables include: density of objects in the display, organization of the objects, meaningfulness of the objects, area of the field of apprehension, rapidity of presentation, mental set of responder, and many others.

This short digression examined a few aspects of a stream of inquiry regarding apprehension of the number property of a set. It seemed appropriate in light of the nature of the rational counting tasks used in the first-grade interviews, and in light of some of the children's performance on the tasks. Examination of the tasks and observation of children's performance on the tasks now continues.

The next rational counting task involved presenting a child with a set of three dot cards; one with 2 dots, one with 3 dots, and one with 4 dots. They were instructed to indicate the card with three dots. The same procedure was used with six dots (with a 5-dot and 7-dot card in the display), ten dots (with a 9-dot and 11-dot card in the display), and thirteen dots (with a 12-dot and 14-dot card in the display).

As you would expect more errors occurred in choosing the greater number cards than the lesser number cards. Generally observations by (E) suggested three broad categories of techniques used by children in responding to the tasks. There were the serious counters, the more impulsive global choosers, and those that seemed to choose the means to fit the task and combined the use of both methods. Generally these observations confirmed the various categories of counting techniques reported by Denmark (1975).

The serious counters counted every card, even those with two, three, and four dots. Even if the requested card were the first counted, they would count the other two cards in the trio presented. As observed by Denmark (1975), there were both visual counters (no pointing to and/or touching each dot) and counting that incorporated overt motor responses of the fingers pointing to and/or touching each dot. The "memory component" appeared to enter into the task. Although the enumeration of a given card was an exhaustive one, and therefore they did not have to recall where to stop counting, they would often have to be reminded of the number they were looking for. Errors of the serious counters generally occurred with the greater numbers and often involved the component of counting examined by Potter and Levy (1968), the ability to point to (or look at) each item in the array, one at a time, until all had been taken exactly once. Errors would occur when a dot was skipped (omission) in the process, or would be counted more than once (redundancy). As with both Potter and Levy (1968) and Denmark (1975) there appeared to be students who systematically ordered the spacial field in enumerating it, while others would use (what appeared as) a rather random enumeration of the field. Insufficient evidence was available on which processing procedure (if either) was more closely associated with errors in enumeration. Some (E) commented, however, on the rather incredible ability of certain "random enumerators" to be accurate in their counts and continuously avoid omissions and redundancy.

The more impulsive choosers would generally react very quickly to a direction and choose a dot card. There was little evidence of counting. The observation was that this procedure would often work quite well for the 3-dot and 6-dot tasks but led to errors on the 10-dot and 13-dot tasks. The digression a few paragraphs back to studies that examined apprehending ability may be apropos to these observations. Recall the hypotheses regarding two different mechanisms for judging numbers, one for numbers of objects of about six or less, and a different process for numbers greater than six. It appeared that the success some students had in quickly identifying the 3-dot and 6-dot cards (and the positive reinforcement for that success from E) created a mental set that compelled some to continue responding impulsively. They seemed unaware that a shift in processing procedure may be necessary as the numbers got greater. For the majority of students the task with 10-dots and 13-dots poses a "problem" since it is generally beyond their ability to give a quick and accurate habituated response. Luria (Callahan, 1975) has suggested that an important phase of problem solving involves the restraining of impulsive responses and the investigation of the conditions of the problem. This was quite obviously absent from the impulsive chooser's reaction to the tasks with larger numbers.

On the other hand, there were some students who seemed to be aware that they could react impulsively at the 3-dot and 6-dot levels, but needed to switch processing procedures with the 10-dot and 13-dot tasks. Some few students actually verbalized this to (E). After choosing the 3-dot and 6-dot cards quickly they said, "Now I have to count." They seemed to be able to restrain an impulsive response and be aware of the different conditions of the task.

The next rational counting task involved sorting 3-dot, 6-dot, and 10-dot cards into appropriate piles. Three criterion cards (one with 3-dots, one with 6-dots, one with 10-dots) were attached to a folder. A deck of 12 cards (four with 3 dots, four with 6 dots, four with 10 dots) was shuffled and the task demonstrated by (E). The child was then handed the remaining cards to place in appropriate piles. The arrangement of the dots on each 6-dot card, for example, was different. There were no cards in the deck that did not belong; it was an exhaustive classification of the three numbers.

Generally this task was somewhat easier than the previous task. Again there were the serious counters, the impulsive sorters, and those who would use both procedures depending on the number of dots involved. The "memory component" was again operational with some of the counters. They would count the top card in the deck (example, six); by then they would have forgotten the number of dots on the three criterion cards and would have to count those again before the choice could be made. In some severe cases this ping-ponging of forgetting the number of dots on the sort card and the criterion cards happened continuously during the task. It would appear that the impulsive sorters fared better on this task. The discrimination was much grosser and the sorting of 3-6-10 dot cards was probably much less a "problem" task. The impulsive sorters did have some problems when they were faced with similar forms of arrangements of dots on a card though the numbers were different. For example, if the two cards below would appear consecutively in the deck to be sorted it would often lead to error by the impulsive sorters since they would place them in the same pile. Such

similar forms seemed to have less effect on the counters. Some of the students would impulsively sort the 3-dot cards and 6-dot cards and then count (at least one or two) of the ten-dot cards.

The final tasks involved ordering a set of seven dot cards (one through seven dots inclusive) from least to greatest; and a set of seven numeral cards (one through seven inclusive) from least to greatest. A general observation was that when there was familiarity with the numerals, that was an easier task than ordering the dot cards. In ordering the dot cards all the behaviors discussed in the previous tasks manifested themselves. Some counted. Some did not. Some would impulsively order through about five and then would have to count to determine which of the two remaining was the six-card and seven-card. In ordering the numeral cards it was not unusual to have the numerals backward or upside down in the sequence, but still have them ordered correctly. Again, the "memory component" would often come into play in both tasks. While examining the next card to be ordered, some would forget how far they had already ordered the cards and would go back to the beginning to count up to where they were in the sequence.

These then were the tasks on the first grade interviews that involved counting.

An opportunity arose to relate the counting performance of the approximately forty first-graders described in the previous paragraphs with their performance on an arithmetic achievement test. Subsets of tasks on the standard achievement test were examined. On the "Concepts" part of the test, two subsets were examined tasks which did not require counting (Example: "Mark under the road that is the widest") and tasks which probably involved counting ("Mark under the group that has the most pieces of candy.") On the "Problem Solving" part of the test, those items where addition was required ("Paul has three cents. His mother gives him three more. How many cents does he have then?") were clustered and those requiring subtraction ("Barbara had six jacks. She lost two of them. How many does she have now?") were clustered. Likewise, on the "Computation" section the addition (2 + 6 = __) task and subtraction (4 - 3 = __) were considered separately. Three aspects of the counting tasks in the interview were examined: rote counting forward, rote counting backward, and rational counting.

A given student's performance on two tasks (one counting, the other a test task) was entered in a table. Illustration 1 indicates an entry, S, for a student who rote counted to 19 and performed at a difficulty level of 0.48 (number of items correct/total number of items) on the "concept" items requiring no counting.

ILLUSTRATION 1—PAGE 90

A median-split was then used to "collapse" the data in a 2 x 2 table. Frequency of students above the median (A) on one set of counting tasks and above the median (A) on a set of test tasks; above the median (A) on one set of counting tasks and below the median (B) on a set of test tasks; below the median (B) on one set of counting tasks and above the median (A) on a set of test tasks; and below the median (B) on one set of counting tasks and below the median (B) on a set of test tasks, are shown collectively in Table 1.

TABLE 1--PAGE 91

The frequencies within a median row or column of the grouped data are not reflected in the 2 x 2 tables, therefore the fluctuations in the total number of students reflected in the tables. They relate and reflect the more extreme performances on the test tasks and counting tasks. A great deal of caution must be taken in drawing relationship conclusions from such an unsophisticated analysis. It was done to simply get a gross feeling for some possible clues and cues for more penetrating analysis of data gathered in the future.

Some observations on rote counting forward and test tasks:
- a positive correlation could probably be expected between rote counting performance and performance on achievement test tasks.

- the deviant, BA-AB, cells suggest that there may generally be more students below the median on rote counting performance but above the median on test tasks than vice versa.

- rote counting ability may be generally related to achievement test item performance rather than specifically and differentially related to various subsets of items.

Some observations on backward counting and test tasks:
- a positive correlation could probably be expected between backward counting performance on achievement test tasks. This correlation might be greater in degree than the correlation of forward counting and rational counting with achievement test tasks.

- backward counting ability may be generally related to achievement test item performance rather than specifically and differentially related to various subsets of items.

Some observations on rational counting ability and test tasks:
- a positive correlation could probably be expected between rational counting performance and performance on achievement test tasks.

- the deviant, BA-AB, cells suggest that there may generally be more students above the median on rational counting performance and below the median on achievement test performance than vice versa.

- rational counting ability may be generally related to achievement test item performance rather than specifically and differentially related to various subsets of items.

It would appear that performance on the various counting tasks contained within the first grade interview are positively correlated with performance on a standardized achievement test. Although the PMDC Mathematics Test-Grade One contains many items other than those associated with counting performance, they report correlations of 0.79 between the PMDC and the Key Math measures, and 0.72 between PMDC and the Otis-Lennon (1975). The initial crude analysis carried out here would suggest a rather high degree of correlation may exist between the ability to perform the various counting tasks and performance on a standardized achievement test. The relationship would appear to be general rather than specific to certain subsets of achievement test tasks.

The students in the small group reported on in this paper were first graders in two different classrooms two years ago. They are presently third graders. I recently attempted to follow-up on how they are performing in school arithmetic. Of the approximately 40 students, there are twenty-one still in the school at the third grade level. Some have been held back a grade, others have evidently moved out of the school. Of the twenty-one, twelve performed at or below the fourth stanine on a standardized test at the end of the second grade. Another administration of a standardized test at the beginning of third grade confirmed this rather low performance on standardized tests of arithmetic. The nine others in the third grade were functioning above the 4th stanine on a standardized test at the end of second grade. Table 2 presents a comparison of performances of the underachievers (lower 4 stanine) with the achievers (above the 4th stanine).

TABLE 2--PAGE 92

Although much caution must be used in making inferences from such a small and non-random group of children, there is at least a suggestion that the counting tasks have some degree of predictive validity in regard to arithmetic achievement.

This paper started by presenting some authoritative comments on the importance of counting as a basic skill of arithmetic. Evidence from some "pilot" administrations of counting tasks to first graders suggested a relationship between counting tasks and standard achievement test performance. Finally, for a small number of students involved with the pilot administration, it would appear as if skill in counting on entering first grade may have some predictive value for success in school arithmetic. These observations must be very tentative because of small numbers and lack of rigor in design and analysis.

Observing young children involved with performing simple counting tasks makes one aware of the complexity of these deceptively simple tasks. In microchosm they make demands on many of the complex cognitive processes needed for success in school arithmetic. Counting tasks, although reacted to differently by different students, seem to incorporate in some ways the entire range of synthetic mental activities mentioned by Luria (1973): perception, movement and action, attention, memory, speech, and thinking.

In speculating on some implications regarding counting in the school program, it may be useful to take a further look at some of the work of Luria in the area of neuropsychology. He presents three basic laws governing the work structure of the individual cortical regions of the brain. The first is the law of the hierarchical structure of the cortical zones: the primary zone, the secondary zone, and the tertiary zone. This system does not remain the same but changes in the course of ontozenetic development. The second is the law of diminishing specificity of the hierarchically arranged cortical zones: the primary zones possess maximal modal specificity, in the secondary the specificity is present to a lesser degree, in the tertiary zones there is even less specificity. The secondary and tertiary cortical zones have a predominance of multi-modal and associative neurons. There appears to be progressive transfer from the primary regions to the secondary and tertiary where specialization of function seems to occur.

By the time students enter school, development from the primary regions to the secondary and tertiary should be well underway. Continued appropriate development, then, would seem to be best nurtured by an environment rich with opportunities to relate, to associate, to develop voluntary attention within complex settings, and to develop the higher cognitive functions.

A simplistic reaction to the role of counting in achievement may be to directly and simplistically work on counting, backward, forward, and rational, outside of a meaningful setting. This would probably do more harm than good. In light of the developmental needs of the child, counting experiences should grow from a socially stimulating setting where a sensitive teacher subtly involves the child in activities where counting takes place within a meaningful setting. Then the dynamic relationships which mark cognitive development will be best served.

One of our jobs is to continue to search for the right questions to ask our students, to observe behavior, to reflect on the meaning of that behavior within the broader contexts of intellectual growth and development, and to firm-up the inference lines between learning and development. Then perhaps we can offer teachers the kind of professional preparation suggested by Mayer (1961) and cited by Lovell,

> What future teachers need, and cannot find, is the course which attempts to explore the profound aspects of the deceptively simple material they are going to teach, which analyzes case by case the types of difficulty that children find in approaching such material...

REFERENCES

Beckwith, M. & Restle, F. Process of enumeration, Psychological Review, 1966, 73, 437-443.

Callahan, L. Neuroscience and mathematics education, Proceedings of the Third National Conference on Remedial Mathematics, J. Heddens, & F. Aquila, (Eds.) Kent State University, Kent, Ohio, 1975.

Conference on Basic Mathematical Skills and Learning, U. S. Department of Health, Education, and Welfare, National Institute of Education, Euclid, Ohio, 1975.

Denmark, T. Diagnosis of entry concepts and skills: Grades one and two, Proceedings of The Third National Conference on Remedial Mathematics, J. Heddens, & F. Aquila, (Eds.) Kent State University, Kent, Ohio 1975.

Gagne, R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1970.

Jevons, W. S. Nature, 1871, 3, 281-282.

Kaufman, E.; Lord, M; Reese, T. & Volkmann, J. The discrimination of visual number, American Journal of Psychology, 1949, 62, 498-525.

Luria, A. R. The working brain. Basic Books, New York, 1973.

Mayer, M. The schools. Garden City, Doubleday Books, 1961.

Ogletree E., Rackauskas, J., Buergin, T. Teaching number sense through rhythmical counting, Elementary School Journal, 1970, 71, 11-17.

Potter, M. & Levy, E. Spatial enumeration without counting, Child Development, 39, 1968, 39, 265-272.

Saltzmann, I. & Garnes, W. Reaction time as a measure of span of attention, Journal of Psychology, 1948, 25, 227-241.

Tanes, E. H. Two mechanisms for the perception of visual numerousness, Archives of Psychology (no. 265) September, 1941.

Wang, M.; Resnick, L.; Boozer, R. "The sequence of development of some early mathematics behaviors, Child Development, 1971, 42, 1767-1778.

Warren, H. C. The reaction time of Counting, The Psychological Review, 1897, 4, 570-591.

TEST CONCEPTS TASKS (Counting)

Rote Counting	00-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
00-10					/					
11-20		/			⑤ /				✗	
21-30						//	//	/	/	//
31-40					/		///	/	/	
41-50								/		
51-60					///					
61-70										
71-80										
81-90										
91-100					//	//	/	///	///	//// /

Illustration 1. Median Split Procedure

Table 1

Frequency above (A) and below (B)
Median on Counting and Test Tasks

TEST TASKS

		Concepts				Problem Solving				Computation			
		No Counting		Probably Counting		Addition		Subtraction		Addition		Subtraction	
		B	A	B	A	B	A	B	A	B	A	B	A
Rote Counting	A	5	9	4	12	5	11	5	11	5	10	4	8
	B	9	7	7	7	9	10	9	9	7	7	10	8
Backward Counting	A	6	12	2	17	6	15	5	16	8	10	5	12
	B	9	4	12	4	13	5	14	4	13	4	11	5
Rational Counting	A	6	6	6	9	7	10	5	12	9	8	8	7
	B	7	2	6	4	7	5	8	3	6	2	6	2

Table 2

Comparison of Counting Skills at first-grade Entrance
for Two Groups of Third Graders

	N	Counted to	N	Counted Backward from	N	Demonstrated with Blocks	N	Discriminate Dot Cards	N	Sorted Dot Cards	N	Ordered Cards
ow	3	39	1	6 & 10	3	4,8,12,16	2	3,6,10,12	3	3,6, & 10	2	both arrays
chievers	2	29	1	6	3	4,8,12	3	3,6,10	1	3,6		& numerals
	2	27	10	Unable to	2	4,8	2	3,6	8	None	4	arrays only
	2	20		count	4	4	2	3			6	neither
	1	15		backward			3	None				
igher	1	100	1	6, 10, & 13	3	4,8,12,16	6	3,6,10,12	8	3,6, & 10	5	both arrays
chievers	2	59	4	6 & 10	6	4,8,12	1	3,6,10	1	None		& numerals
	2	39	1	6			2	3,6				
	3	29	3	Unable to							4	neither
	1	19		count								
				backward								

Reaction Paper
to
Counting Performance and Achievement:
Some Preliminary Observations

James W. Heddens
Kent State University

The three quotations that Dr. Callahan has cited in his paper seem to place counting in three different perspectives. Buchanan, in his statement, places counting as the primary base upon which a mathematics program is developed. Catherine Stern, in her book Help Children Discover Arithmetic, seems to place emphasis upon relationship first, followed by the development of number, then the development of numeral. Piaget's work seems to place counting in a different perspective.

Hilton's quotation places his emphasis upon competence in the use of the number system. One does not know if counting, comparing, ordering, and measuring are placed in a significant order. Piaget would consider comparing prior to counting. He would follow comparing with rote counting, then rational counting. If a child has considered comparing, then moves to counting, is he using rote counting or rational counting?

Rubenstein commences his statement with the word "understanding" and as soon as this is indicated then we are considering rational counting and not rote counting. He has sequenced counting first, followed by comparing, then ordering. Again, this seems to disagree with Piaget's work.

Callahan has used excellent procedure in defining the difference between rote and rational counting. He has done this by suggesting different types of tasks for each type of counting. The definitions would be more meaningful for me if some type of evidence could be presented that would verify that the tasks really do indicate either rote or rational counting. In my opinion the task of asking a child to count forward and the task of asking a child to count backwards are two different levels of functioning. I have difficulty placing both of these tasks under the heading of rote counting.

Gagne's theory of verbal chaining seems to be an excellent technique for evaluating the level of functioning in rote counting. Then we must assume that the longer the verbal chain, the higher the level upon which the child is functioning.

Callahan's observation that if a student could count to sixty then he could make it to one hundred seems to indicate more than rote counting. Evidently, the student was able to generalize a pattern and apply that generalization in order to count to one hundred. Has the child moved from rote counting to rational counting? Then the question needs to be considered at what point in the child's development are we willing to say a child changes from rote counting to rational counting?

The linguistic explanation suggested by Callahan seems appropriate. The linguistic work of Hargis seems to support Callahan's analysis. I feel that more research from the mathematical and linguistical approaches needs to be considered.

Callahan's observation that counting to 100 was more enjoyable for young children seems to me to be a logical conclusion. Parents seem to prize counting and often ask children to count. The younger the child, the more positive reward is obtained by counting further. Positive strokes will encourage the child to concentrate on counting farther. At what point will the child discover a numeration pattern and switch from rote counting to rational counting? Maybe parents receive more self gratification from their child's counting than the child himself receives. Thus the parent will place more pressure upon the child to count further. Consequently, we have developed a cyclic effect.

Grossnickle and Brueckner in their 1959 edition of Discovering Meanings in Arithmetic have done an excellent job in describing the nature of counting. They define six different stages in the complete process of counting. The six stages are (1) rote counting, (2) enumeration, (3) identification, (4) reproduction, (5) comparison, and (6) grouping.

Rote counting is defined as a mere repetition of numbers in sequential order without meaning. Grossnickle and Brueckner do not indicate how an individual evaluates a child's counting to know if the saying of the number name in sequence has meaning for a child.

Enumeration is defined as rational counting and means to find the number of objects in a set. Thus a child must be able to use one-to-one correspondence in counting the objects in a set. It would seem to me that in observing children the way the child functions would indicate his level of understanding. As I have observed young children, counting is a touching process with the utilizing of one-to-one correspondence. Ask a five year old child to count the number of children in a room and he will proceed around the room and touch each person as he says a number. A next higher level seems to be apparent as the child stands and points to each individual as he counts using one-to-one correspondence. Now he is still using one-to-one correspondence but he has placed a distance between the objects being counted and the finger. The next higher level seems to be when a person nods his head and counts using one-to-one correspondence. I have observed children who use touching their own nose with fingers as they count or clicking their teeth as they count. A more sophisticated level of counting is when a person counts by directing his eye to each member as he counts. This is probably a method that is still being used by adults as they count.

Identification is defined as answering such questions as which set has four spoons. A child could answer the question either by counting or by mere recognition. Maybe we encourage children to use immature methods when we ask a child to count the number of spoons in the set. How do we know if the child could have recognized the fourness without counting? In my opinion we encourage children to depend upon counting to provide responses rather than encouraging them to learn recognition. In our work we use flash cards with dots on them and flash them rapidly and ask the children how many. Most five year old children can handle up to five dots without counting.

Reproduction is defined as giving the correct response to such questions as "Give me four of the spoons." A child must respond by selecting four spoons from a larger set of spoons. I would also observe how the child behaved. Did he count out four spoons, or did he count by twos to get four spoons, or did he merely recognize four spoons and pick them up? Again I see subdivision to the categories, the classification suggested by Grossnickle and Brueckner.

Comparison is utlized when observing a set and answering such questions as "How many more silver spoons are there than plastic spoons?" In answer to that question, it seems to me there are three different ways in which the child can function.

Grouping is shown when the child identifies at a glance the number of objects in a subset without counting and then counts from that point on. The child again has several options as to how he can function.

The research from the early forties done by Carper needs to be replicated as well as the work of Dan Dawson. Maybe it would be to our advantage to reexamine the group test devised by J. Kern at University of Freiburg for testing maturity levels of counting.

Callahan has gone on to utilize more tasks and to discuss children's responses. I feel this is an excellent beginning. Yet many questions keep appearing that should be considered. For instance, should more theoretical analysis of the content be attempted prior to collecting responses from children? After logically sequencing the theoretical levels of counting can tasks be designed that can be validated to verify that each task is indeed functioning on the level specified? I feel the technique used by Nichols and Denmark to be excellent. After a protocol has been established, follow the protocol and videotape each case during response. Thus the researcher does not need to think about the protocol, observe the child's behavior, record the child's behavior and try to interpret the behavior all at one time. Now the researcher can concentrate on the protocol solely, for the videotape is recording the behavior. The researcher can interpret and analyze the tape at a future time when he can zero in on a specific aspect. The tape is also available for multiviewing, thus providing an opportunity for several different individuals to interpret the data and for several viewings of the child's behavior.

A PERSPECTIVE ON THE FUTURE

James W. Heddens, President
Research Council for Diagnostic and Prescriptive Mathematics

One can observe young children having difficulty learning mathematics. We have reading centers for children suffering with reading difficulties, we have specialists to work with children suffering from seeing or hearing difficulties, we have hospitals devoted to helping children with physical difficulties, and many more special areas. Where do we have professional trained individuals working with children suffering from mathematics difficulties such as acalculia or dyscalculia? Why haven't the individuals responsible for mathematics developed and prepared instruments, materials and techniques for "remedial mathematics"?

Over a period of years we have observed Guy Bond working in this area of reading at the University of Minnesota, we have observed Durrell at Boston College, Sheldon at Syracuse, and many others. When we study their work we find much repetition. Why didn't these individuals pool their talents and resources? Why does each individual have to reinvent the wheel? This has bothered me for years, consequently the questions are:

1. Can the mathematics people cooperate in their research?
2. Can diagnostic and prescriptive research in mathematics education be organized nationally or internationally?
3. Can duplication of work and effort be avoided?

With these questions in mind and the cooperation of KEDS General Assistance Center, Kent State University organized the First National Conference on Remedial Mathematics in 1974. As a result of this first conference, papers and reaction papers have been prepared and these now are published and available from ERIC. In my opinion these papers and reaction papers are a major contribution and seem to be setting a bench mark.

In 1975 the Second National Conference on Remedial Mathematics was held at Kent State University. A different format was used. No papers were produced and consequently there is no published material. We made and used some video tapes, but I am not sure that the video tapes are of a quality that we would want to reproduce.

In 1976 the Third National Conference on Remedial Mathematics was held. In conjunction with the Third National Conference we concentrated on research. As a result of this conference it was decided to organize a research council. Bill Speer presented to the group a very complete bibliography and compiled a dictionary of terms focusing on the diagnostic-prescriptive arena.

A steering committee was appointed to lay the foundation for a new organization--The Research Council for Diagnostic and Prescriptive Mathematics. It was decided at this 1976 meeting to have membership open to all interested individuals--mathematics teachers, supervisors, coordinators, mathematics professors, mathematics education professors, and researchers.

Where are we today? Where are we headed? We now are a group with a Constitution and By-laws, elected officers, and a solid membership.

I would like to blue sky with you for a few minutes and project as to what I see as our future. I am not interested in sheer numbers of members. I am interested in sincere, dedicated, humane individuals who are willing to cooperate, to develop new and exciting research, and to help people learn mathematics.

In my opinion the most important position that has been established by our new Constitution is the Vice-President for Research. This position carries with it the task of helping all of us cooperatively organize research in mathematics education throughout the United States and Canada so that a united front will develop from within mathematics education. Can we eliminate duplication? Can we get teams of researchers (such as classroom teachers, supervisors, coordinators, administrators, and professors) working together in research projects to solve our problems, to synthesize innovative approaches, to create insightful diagnostic instruments and diagnostic techniques, and to develop new and interesting material? Can the Vice-President for Research organize this united front through the Research Council for Diagnostic and Prescriptive Mathematics? Can a consortium be created with each unit researching a portion of a much larger problem so that all the pieces fit together into one large puzzle? Can we furnish different samples from various parts of the country? At the Maryland conference we had participants from thirty five different states, nine Canadian provinces and West Germany. What would result if some were willing to locate samples of students to participate in a large research project? If each was willing to use a designated protocol with a sample of children? If each was willing to help collect research data?

I feel we have the dynamic embryo for the most extensive research project ever undertaken and the potential can be found in our members. You are the person-power that is needed. If we are organized, if we can define the needed research, then why can't we develop proposals for federal or private foundation funding? Visualize what an impact we can have on mathematics diagnostic and prescriptive teaching.

I have shared with you some of my ideas and feelings about our future. Can we accomplish such a mammoth undertaking? Anything is possible with a positive attitude, a united front, and hard work.

Do you want to be a part of this dynamic organization? We would enjoy having each of you working as part of a team.

A united front needs to be developed to organize and to avoid duplication of research in this area. Diagnostic instruments need to be developed, techniques of diagnosing need to be developed, and unique methods of corrective teaching need to be developed. Researchers will be needed at all levels--in the clinic and in the classroom, professors, clinicians, supervisors, and classroom teachers. There is room for everyone to actively participate in the Research Council for Diagnostic and Prescriptive Mathematics.