DOCUMENT RESUME

ED 040 536 EC 005 826

AUTHOR Haring, Norris G.; Hayden, Alice H.
TITLE Instructional Improvement: Behavior Modification.
INSTITUTION Child Study and Treatment Center, Fort Steilacoom, Washington.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Elementary and Secondary Education.
PUB DATE 68
NOTE 395p.

EDRS PRICE EDRS Price MF-$1.50 HC-$19.85
DESCRIPTORS Administration, Autism, *Behavior Change, Behavior Development, Behavior Problems, Computer Assisted Instruction, Controlled Environment, *Educational Technology, Evaluation Methods, *Exceptional Child Education, *Experimental Programs, Operant Conditioning, Program Evaluation, Reading Instruction, Research Needs, Special Classes, Student Evaluation, *Teaching Methods

ABSTRACT

Sixteen papers are provided. B. F. Skinner discusses the arrangement of contingencies for learning: Lloyd Homme describes behavioral engineering; and Frank Hewett considers behavior modification in special education. Also treated are experimental education by Norris Haring, program evaluation by Arthur Lumsdaine, and administration of special classes by Harold Kunzelmann. John Cawley presents a system of initial reading instruction; Max Jerman surveys computer assisted instruction; and Thomas Robertson examines the impact of educational technology. Further papers are on teaching children with behavior disorders by Richard Whelan, developing cooperative social behavior by Laurence Peter, providing academic and social classroom management by Harold Kunzelmann, and using operant reinforcement with autistic children by Charles Ferster. In addition, Thomas Lovitt sets forth a basis for systematic replication of a contingency management classroom; Richard Kothera discusses educational environments and administration; and Max Mueller reviews trends in research in the education of the handicapped. (JD)

Workshop in

Instructional Improvement: Behavior Modification

Norris G. Haring, *Director* / Alice H. Hayden, *Coordinator*

CHILD STUDY AND TREATMENT CENTER
FORT STEILACOOM, WASHINGTON—JUNE 17–JULY 17, 1968

Supported through ESEA Title VI-A funds (Project No. 57-27-400-8-602).
University of Washington Course No. Education X474.

Workshop in

INSTRUCTIONAL IMPROVEMENT: BEHAVIOR MODIFICATION

Norris G. Haring
Director

Alice H. Hayden
Coordinator

Child Study and Treatment Center
Fort Steilacoom, Washington
June 17 - July 17, 1968

Supported through ESEA Title VI-A funds
Project Number 57-27-400-8-602
University of Washington Course Number Education X474

CONTENTS

ACKNOWLEDGEMENTS

Assistance and support throughout the planning of the workshop were provided by Helena Adamson, Hy Henderson, Dohn Miller, James B. Piland, and S. L. Shiemo. The actual success of the workshop must be attributed to this group, to the activities of the demonstration teachers (Ralph Bohannon, Betty Casperson, Hal Caufield, Rolene Caufield, Tal Guppy, and Sharon Marks), and to the Coordinator of Special Education and Title VI project for Clover Park Schools, James B. Piland. We wish to thank Pat Leventhal, Penny de Bishop, Pat Davis, and Nancy Burke for their secretarial assistance and Charles R. Brown for his technical assistance with the video and audio instrumentation. Our special gratitude is extended to the administrators of the Clover Park School District: Mr. T. Olai Hageness, Superintendent of Schools and Dr. William G. Kalenius, Director of Pupil Services.

CONTRIBUTORS

John F. Cawley
Associate Professor of Education
University of Connecticut

Charles B. Ferster
Professor, Department of Psychology
Georgetown University

Norris G. Haring
Professor of Education and Director
Alice H. Hayden
Professor of Education and Associate Director

Experimental Education Unit
Child Development and Mental Retardation Center
University of Washington

Frank M. Hewett
Chairman, Special Education
University of California, Los Angeles

Lloyd E. Homme
Director, Research Department
Behavior Systems Division
Westinghouse Learning Corporation

Max Jerman
Research Associate
Institute for Mathematical Studies in the Social Sciences
Stanford University

Harold P. Kunzelmann
Lecturer in Education and Acting Principal
Experimental Education Unit
Child Development and Mental Retardation Center
University of Washington

Richard J. Kothera
Superintendent
Lombard, Illinois
School Number 15

Thomas C. Lovitt
Associate Professor of Education
Experimental Education Unit
Child Development and Mental Retardation Center
University of Washington

Arthur A. Lumsdaine
Professor of Psychology and Education, and Chairman
Department of Psychology
University of Washington

Max Mueller
Chief, Projects and Program Research Branch, Division of Research
Bureau of Education for the Handicapped
U. S. Office of Education

Laurence J. Peter
Associate Professor of Education
University of Southern California

Thomas F. Robertson
RCA Instructional Systems

B. F. Skinner
Edgar Pierce Professor of Psychology
Harvard University

Richard J. Whelan
Associate Professor and Chairman
Special Education
University of Kansas

OVERVIEW OF THE WORKSHOP

Any endeavors to improve methods of teacher instruction must aim toward the improvement of conditions for learning. Increasing the competencies of the classroom teacher, however, cannot be accomplished by devising more complete instructional methods; for methods of instruction are closed systems, unable to change with the variation in patterns of child response. Instead, what is needed is increased emphasis on systematic procedures.

Almost all educational investigators and teachers use observation as a primary tool. But in order to systematically measure and evaluate the interrelated events and behaviors that occur during learning, observation procedures need to be refined. Consequently, the Workshop in Instructional Improvement: Behavior Modification emphasized the need for refinement and use of procedures for direct observation, continuous measurement, and systematic changes in classroom conditions in order to produce effective instruction and learning.

Objectives

As the professional training program defines and requires the implementation of these procedures by the teacher-in-training, the trainee should acquire facility in:

1. Assessing each child's skills in the critical areas of instruction.
2. Arranging instructional cues, based on child performance.
3. Measuring precisely the child's responses to the program.
4. Systematically arranging the environment, including those reinforcing events in the environment that increase performance.

The objective set forth in this institute was consistent with the idea that the goal of any instructional training program for teachers, then, can be achieved through the arrangement of learning experiences which build cumulatively on their competencies and lead toward precision in instruction. Specifically, the institute was designed to increase teacher competencies in:

1. Program development and analysis.
2. Refinement of instructional procedures.
3. Assessment of pupils' academic and social skills.
4. Recording and graphing of pupils' performance and behavior.
5. Application of reinforcement principles for motivating pupil performance.

These training objectives were accomplished through three kinds of experiences: (a) didactic, (b) demonstration and observation, and (c) application of instructional procedures under direct, continual supervision.

Training Program

The children who served in the workshop, selected because of the mild to severe learning and behavior problems they exhibited, came primarily from the Clover Park Schools. One of the demonstration classrooms comprised children regularly enrolled at the Child Study and Treatment Center of Western Washington State Hospital, who exhibited a wide variety and degree of behavior disorders.

During the four-week period, the participants spent two and one-half hours each morning with pupils. At first the demonstration teachers functioned in the classroom while the participants observed or assisted. As the institute progressed, the participants became the classroom teachers, under the direction of the project teachers. Afternoon sessions consisted of lecture-demonstrations and question and answer periods led by the guest speakers.

The first two weeks of the workshop were devoted to didactic experiences, observation, and demonstration. During this period, teachers-in-training gained experience with specification, measurement, and recording and graphing of responses. Opportunities were provided to assess the academic skills of pupils and to select instructional programs. Analysis of the child's responses to instructional programs provided the basis for the development of instructional sequences.

The second half of the workshop concentrated on specific demonstrations of instructional programs, display instrumentation, and reinforcement procedures. Students were afforded opportunities to imitate demonstration procedures and to apply their new skills to a wide variety of instructional problems.

In keeping with the objectives set forth for the workshop, specialists from all parts of the United States were invited to share their knowledge with the participants. The individuals selected were those who had demonstrated outstanding achievements in the application and investigation of behavior variables.

It seems only fitting that the introduction to these Proceedings should be taken from the presentation by B. F. Skinner, who more than any other individual has been instrumental in the development and refinement of the principles and procedures basic to the workshop.

The area of programmed instruction was strongly represented in the speaker series. Arthur A. Lumsdaine gave the introductory comments in his lecture, "Assessing the Effectiveness of Instructional Programming." This was followed by a lecture on the variables of initial reading instruction by John F. Cawley. Then Max Jerman spoke on "Computer-Assisted Instruction."

To improve classroom instruction, pinpointing and measurement of behavior must become increasingly more precise. Information concerning systematic procedures for response measurement was presented in two lectures by Harold P. Kunzelmann.

After this exposure to programming principles, programming in the content areas, and response measurement, the participants attended discussions on the importance of systematic contingencies of reinforcement for motivating the learner. Charles Ferster described the procedures and effects of operant reinforcement in infantile autism. Thomas Lovitt was selected to present a practical position on a contingency management classroom. Then, turning to emerging views concerning the many ways in which behavior modification can be achieved in the classroom, four well-known behaviorists were invited to present their versions of the practical applica- n of behavior modification. They were: Frank M. Hewett, Lloyd E. Homme, Laurence J. Peter, and Richard J. Whelan.

The importance of the administrative staff and the general educational environment cannot be underestimated, since these also influence performance. For this reason, Richard J. Kothera, a long-time school administrator, was invited to share his ideas on administration.

The continuing impact of computer technology on education is reaching the point where information about computer applications in the classroom and in education must be recognized. Thomas F. Robertson, in his presentation, "The New Impact: Technology in Education," performed an important service in th. area.

Finally, with a presentation on recent research trends in teaching handicapped children, Max Mueller summarized the lecture series of the workshop.

Evaluation

Each participant in the six groups was evaluated by the demonstration teacher to whose class he was assigned on the basis of a project conducted during the practicum. In most instances, this project included the requirement that responses be pinpointed, recorded. and plotted as rate on six-cycle log paper.

A follow-up evaluation was conducted during the Winter of 1969 in the classrooms of the workshop participants. Workshop staff observers were especially interested in observing the carryover, in both quantity and quality, of the various types of procedures learned during the summer. The observers looked specifically for similarity in instructional procedures, instructional materials, and simple instrumentation introduced during the workshop.

Follow-up evaluations in the classrooms of the participants were also conducted by the four demonstration teachers. The four resource teachers were particularly interested in observing the use of four different features of systematic instruction. First, they made note of the use of programmed materials or the adaptation of regular classroom materials to permit active responding and response measurement. They also looked for procedures of response measurement, including the use of event records. Third, they were interested in the types of reinforcement procedures used. Fourth, they noted the degree of contingency management functioning. One further basis for evaluation was the degree to which each participant had carried out his original plans made at the end of the summer workshop.

The project teachers found that procedures taught during the workshop resulted in a broad range of applications by participants. One-third of the participants had managed to incorporate all the procedures into their classroom for the major part of the day. Half of the participants were using some of the procedures, usually some form of response measurement and some degree of systematic reinforcement. The teachers in this group more often used reinforcement without response measurement rather than both together. However, this establishment of a contingency management system was the workshop procedure least often carried over to the classroom. Several teachers expressed the difficulty in obtaining adequate materials. Only a few of the participants had failed to institute any of the procedures. Almost all participants expressed interest in attending another workshop which would emphasize the application of these procedures in many different settings.

Participation and Organization

Approximately 37 children were served in the summer program. Thirty teachers and approximately eight administrators participated. Staff personnel included one administrator from the University of Washington and one from the Clover Park School District, five consultants, four demonstration teachers, and one program specialist. In addition, there were 17 nationally recognized speakers who presented lectures and demonstrations.

The request for the workshop was initiated by the Clover Park Public Schools and the Child Study and Treatment Center at Fort Steilacoom, Washington. Application forms for admission to the workshop were distributed and returned to Mrs. Helena Adamson of the State Department of Public Instruction. The selection committee which reviewed the applications attempted to obtain representation in the workshop from schools in all parts of the state and from various groups such as teachers, administrators, psychologists, and counselors. Some districts which wished to initiate or strengthen specific programs requested the admission of teams of varying combinations of professional personnel. Many people assisted in planning the workshop, coordinating its activities, and evaluating its accomplishment of objectives. Through the cooperation of representatives from the State Department of Public Instruction, school districts, institutions of higher learning, and the Child Study and Treatment Center, it was possible to provide a valuable experience for all those who participated in the workshop sessions.

Norris G. Haring
Director of the Workshop

Alice H. Hayden
Coordinator of the Workshop

CHAPTER I TEACHING: THE ARRANGEMENT OF CONTINGENCIES UNDER WHICH SOMETHING IS TAUGHT

Digest of remarks by B. F. Skinner

Much recent knowledge acquired from an experimental laboratory analysis of behavior is having a dramatic effect on instruction. This knowledge is being carried to the classroom in the form of programmed materials and, more recently and more dramatically, in the design of classroom contingencies of reinforcement. Classroom application of experimental analysis is demonstrating that instructional programming can insure that students learn efficiently. The general tenor of this change in educational method is a move from essentially aversive practices to the use of positive reinforcement.

Whether educators like it or not, most students today study to avoid the consequences of not studying; and herein lies a major problem. They play truant or drop out and escape the whole system. Some turn against it aggressively as young vandals or later as those very costly, more mature vandals who refuse to support education when they are in a position to do so.

The techniques of producing--modifying--behavior through positive reinforcement are extensive, as is the degree of relevant technical knowledge. Solutions to effective practical applications are not so extensive, however, although the contributions observable in the procedures at the Child Study and Treatment Center at Western State Hospital, Fort Steilacoom, Washington and the Experimental Education Unit at the

Experimental Education Unit at the University of Washington are making a real contribution toward solutions. Practical application is hindered by a most critical problem--finding events to function as positive reinforcers to the student in order to shape or maintain behavior.

Reinforcing Events in the Classroom

The idea of bringing real life into the classroom, using only those things which are really significant--that is, significant consequences of the behavior being taught--was a move away from punitive techniques in the right direction. Unfortunately this is not feasible. No one can bring into the classroom the real reasons why something is learned. The classroom isn't big enough to hold real life. Furthermore, if educators wait until the genuine consequences of learning begin to take hold, it is too late to have any very great effect on the student.

Contrived reinforcers are necessary to initiate the behavior. Then natural reinforcers will take over and contrived ones can be dropped. The classic research by Wolf, Risley, and Mees (1964), who taught a child born blind with cataracts to wear glasses, exemplifies the importance of contrived reinforcers. They reinforced the child for putting on and wearing his glasses, using bites of food after making him hungry. It is absurd to say that the child still wears glasses to get food; but it was necessary to produce the initial pattern of behavior while the real consequences of wearing glasses--better sight--began to take effect. The real consequence would not produce the initial behavior, but by using a perfectly spurious reinforcer, namely food, for wearing glasses, the behavior resulted and the real consequence maintained it.

What to use in the classroom as reinforcing events remains one of the big technical problems. Depriving all children of food until they arrive at school and then reinforcing their correct responses with bites of food would work, but the inherent problems make this procedure feasible only for very extreme behavioral difficulties. However, teachers could take advantage of this basic biological reinforcer by using a system of tokens or points, exchangeable for delicious desserts at lunch time, as reinforcers for correct responses to academic tasks.

Success as a reinforcer. Some whopping new reinforcer is not the critical addition to the teaching act. The usual reinforcers, if used well and frequently enough, will be effective. The point of a good program is not to lead the student to obtain a very large, novel, or powerful reinforcer, but to provide reinforcement many times through his being successful again and again. The human organism, fortunately for us all, is reinforced just by being successful. Consequently, if material is designed to facilitate correct responses, the resulting frequent success is enough reinforcement for most persons. Not only will the child's behavior change as he learns to do things he couldn't do before, but he will become highly motivated, his morale will improve, and his attitude toward teachers will change. At this point the major effect has been achieved. Most probably teaching can be so designed that a child is reinforced simply and primarily by being successful, but this calls for precise and expert arrangements of conditions.

In the long run, of course, what is learned becomes valuable. A child who learns to write will eventually write his name many places and write letters to his friends, for example. These real effects cannot be used to teach writing to begin with, but they will take over when just "being successful" in printing a letter properly becomes part of the ultimate and natural reinforcing events.

Contrived reinforcers. If more powerful reinforcers are necessary, then they must be obtained and used. Where "being successful" has not become a reinforcing event, contrived reinforcers are necessary. For example, in problem areas such as city ghettos, it is not impossible that students will be paid for learning, not paid for coming to school but paid for making right answers. If a situation is set up in which no cheating is possible and students are paid a certain amount per correct answer, dramatic results in skill development are highly probably. Financing this change in programming can be accomplished with the money normally funneled into these areas. Rather than provide money non-contingently to families in the form of welfare, students could earn money in school based on good performance.

Problems that arise over who gets paid and how much at what stage of response complexity can be guided by the fundamental principle that reinforcers are used only in the quantity sufficient and necessary to establish the behavior pattern desired. The influence of money on behavior has been well demonstrated and is available to educators, if they want to use it, when necessary. Real life reinforcers do come to influence behavior in time. They may come late but they are the only reasons for getting an education. It would be absurd to teach unless there were some natural consequences. Education must have some good reasons behind it and the behavior of the student should as soon as possible be taken over by these good reasons. At the stage when natural consequences predictably influence behavior, spurious, contrived reasons are no longer necessary.

Classroom Management

Once the reinforcers have been determined, they must be related to the behavior to be produced, in such a way that the behavior predictably results. The Child Study and Treatment Center and the Experimental Education Unit, as well as other school systems, are currently facing this specific problem. The directions being taken, although crude compared to results five or ten years hence, are the beginnings of the analysis of classroom management procedures and as such already exhibit a number of exciting techniques.

Establishing contingencies. Determining how to manage a classroom, that is mapping out the contingencies of reinforcement which will bring about changes in behavior most expeditiously, requires that many questions be answered. For example, how is the student's behavior to be sampled? What reinforcer should be used? Would tokens be the most effective? Will a point credit system be used? To what extent and when can these credit points be made exchangeable just for approval from a friend, the teacher, or the student himself?

If the child is to take satisfaction in what he is doing, then a consequence that is satisfying must follow very closely in time the child's success at a given moment. To depend on the child's ultimately feeling that he has done well is settling for a weak effect. The whole point of operant conditioning is to make the reinforcing event immediately contingent upon the behavior that is to be shaped. The question of determining whether to use points, or tokens, or commendations is only a sub-issue. For when the teacher becomes skillful in timing the presentation of reinforcing events for behavior being shaped, almost magical results follow. The student's behavior does change. His attitude toward what he is doing changes and the basic problem has been solved.

Response counting. Because response counting in the classroom is essential but not easy, several tactics must be initiated. First, the move toward time sampling must be made as quickly as possible. Second, the student must be directed toward keeping his own record of behavior and toward recognizing the connections (contingencies) between what he is doing and the consequences for it. Recognition of contingencies in effect opens the door to the influence of automatic reinforcement, thus freeing the teacher to attend to the management of other contingencies for that child as well as for other children.

Programmed instruction. The value of a program of material--programmed instruction--is just beginning to be understood as something which maximizes the frequency with which the student is correct. Much refinement remains to be accomplished in this area, too. Good programs have been written and have, at times, proven quite dramatic. But people who write programs are seldom aware of what can be done with prompting and probing techniques.

Summary of Progress in Classroom Management

The present degree of achievement in a) the development of programmed instruction, b) the establishment of systematic reinforcement contingencies in the classroom, and c) the management of the student in the classroom to produce effects important to his education are already quite dramatic and portend a bright future. This is especially true as the principles, currently being applied very superficially, are the surface of a basic underlying science. For example, the laboratory study of operant behavior has not been standing still. The principles being used now in education were derived from laboratory experimentation of fifteen years ago. In those fifteen years the science itself has moved forward. Educators applying reinforcement procedures to children in classrooms today lag far behind the application of very subtle contingencies of reinforcement which are daily experience in the operant laboratory.

Solving the basic problems in arranging contingencies of reinforcement in the classroom will be an engineering breakthrough that will provide the educators with an enormously increased capability for applying a great deal of knowledge not now being used. Those few principles now applied in education and psycho-therapy are just a sampling of those available for application. Solving this basic problem will have to be accomplished outside the laboratory, however. The application of these techniques to education must be worked out in the actual act of teaching with real students in real school situations.

Progress toward more precise contingency management is not only hindered by the need for an engineering breakthrough but also by the attitude of several groups of educators. The field of education includes a number of professionals who espouse attitudes opposed to instructional planning and classroom management designed according to basic scientific principles. One such attitude appears to rationalize failure by denying the need for the curriculum which the child failed to learn.

Statements which support a) the need for a de-emphasis on book reading rather than a solid reading program for ghetto children, b) the need for students to acquire a sense of meaning of history rather than learning a substantial body of factual material, and c) the need for experiencing excitement in mathematical discovery rather than acquiring a basic body of cultural knowledge are rationalizations for the failure of modern methods to impart any substantial portion of what is already known. However, transmission of a culture is the main goal of education. Furthermore, a good program teaches painlessly and efficiently what somebody else already knows; a good program will teach in one or two hours what might normally take a whole school day, thus greatly freeing the teacher to engage in the important personal relationships between students and teachers, rather than perform as a flesh and blood teaching machine.

Another attitude opposed to scientific classroom management claims that control in teaching threatens the student's individuality and his right to take credit for his own accomplishments. This misconstrued argument apparently states that good teaching is bad, as the teacher receives credit for student knowledge, and, conversely, bad teaching is good, for learning under bad teaching is a credit to the student himself. Unfortunately, it appears from this attitude that one cannot be credited with being good if the environment makes one good automatically. The same argument is presented against designing a culture where people behave well toward each other. Only when everything conspires against the person in such a way that he naturally would behave badly, but nevertheless behaves well, can he be credited with being good.

The real issue here is not whether behavior should be controlled but whether or not it is to be controlled well. That everyone is controlling everybody else all the time must not be overlooked. While powerful teaching in the wrong hands is of concern, it must not negate techniques to improve teaching in the most effective ways. Furthermore, the issue of giving credit for student accomplishment under an ideal school system will become relatively unimportant, for the students will acquire the abilities, skills, creativity, and originality deemed important, regardless.

Reference

Wolf, M. M., Risley, T. R. and Mees, H. Application of operant conditioning procedures to the behavior problems of an autistic child. Behavior Research and Therapy, 1964, 1, 305-312.

CHAPTER II DIMENSIONS OF EXPERIMENTAL EDUCATION

Norris G. Haring

In order to build a basic body of knowledge on the variables of the teaching-learning act, education has an urgent need to become more scientific. This urgency for a scientific approach is felt in all areas of research, training, and service for children and for the profession.

In any field, the history of scientific development is a history of that field refining its procedures for assessing the relationships between independent and dependent variables. This progress has always involved sharpening procedures of direct observation, avoiding introspective and inferential judgments wherever possible, establishing observable control over variables, and experimental (systematic) manipulation of one independent variable at a time. Substantial advancements in education are at the threshold if educators will use these procedures for evaluating the interrelationships between the independent classroom variables of learning and the dependent variables of performance.

To see that children learn requires effective program planning in the classroom, effective training in the colleges, and in-service training in the schools. It is, therefore, the educator's responsibility to insure that our knowledge of the lawfulness of behavior is extended to the classroom and to professional preparation. Otherwise the door remains open to haphazard learning experiences and concomitant deficits in skill development.

Typically, when the educator meets the almost overwhelming task of rearranging the environment to establish conditions for more effective instruction, solutions seem not only difficult to envision, but probably impossible to complete in one effective step. To complicate the problem further, multitudes of teachers have been taught to approach behavior through a causal frame of reference, making the task of modifying target behaviors difficult if not impossible. The teacher, however, holds within her classroom all the power to change poor performance patterns and inappropriate behaviors to acceptable levels.

Experimental education, an end-product of the concern for a more scientific approach, offers educators the guidelines and procedures to improve child performance through more effective classroom instruction. Experimental education might be viewed from at least four dimensions. First is a set of objectives. Secondly, it has as its basis scientific research delineated as principles of instruction and principles of behavior. Thirdly, it presents a set of procedures for classroom instruction and performance measurement. Finally, it provides the opportunity to extend the scientific base of education through the systematic use of common procedures which incorporate known instructional and behavior principles.

The overall objective of experimental education is to improve instructional procedures in the classroom in order to improve child performance. This objective extends from service to children in the classroom, to the professional training of teachers. It is the investigation and application of principles of instruction and principles of behavior through the utilization of procedures of experimental analysis that is the essence of experimental education.

As far back as 1943, Warren identified the stimuli (conditions and events) within the child's immediate environment as a strong influence on the behavior of the child. Since that time, research has well demonstrated the dramatic changes which occur in a child's performance when all the relevant conditions in his immediate environment are recognized and systematically presented according to plan. If a relationship is established between the child's behavior and any conditions for learning within the immediate classroom environment, the educator can readily observe its influence on the child's performance. A simple change in the pattern of the presentation of an independent variable will change the pattern of performance of the child.

Classroom conditions can be arranged to affect classroom performance and behavior in a number of predictable ways. Scientific research has demonstrated a number of effects, some of which have been delineated by Wallen and Travers (1963) as principles of instruction and others by Skinner (1953) as principles of behavior.

Principles of Instruction

Performance is predictably influenced by a number of arrangements of classroom conditions or events which the educator must recognize in planning for the use and evaluation of other conditions for learning. The delineation of principles of instruction which can be seen to influence classroom behavior includes:

1. Events subsequent to the responses being measured function as independent variables of behavior and performance, just as the events antecedent to them do.
2. When cues are introduced to optimize performance, motivation is at its maximum.
3. Practice in applying a principle to a new problem facilitates transfer of training.
4. The child can only respond correctly to a task at his level of skills.
5. Practicing the response to be learned establishes the skill most efficiently.

Consequently, the ideal environment, where the truest assessment of variables is possible, is one where:

1. Design plans include specifications for control of subsequent events, which may function as reinforcement variables.
2. Each child is motivated to work at maximum performance under the conditions being evaluated for instruction.
3. Response requirements of the task are at the appropriate level for each child.
4. Variables introduced for evaluation which allow measurement of responses exactly like those to be required by the actual classroom task.

Principles c. Behavior

Principles of behavior are statements of the lawfulness of behavior observed under specific conditions. These principles, involving quantitative relationships between stimuli in the environment and behavior, define the types of influence that specific environmental events have on behavior as these events occur in a particular time order to the behavior. It may be that these principles, researched by Skinner as well as a host of followers, have been identified by Wallen and Travers under their first two principles of instruction. Principles of behavior, however, require much further delineation in order to specify fully the effects of variables of reinforcement on the acquisition and maintenance of behavior.

Stimuli in the environment gain control over the functioning of the individual in a number of ways. Some responses initially are conditioned by the presentation of a stimulus without any recourse to reinforcing events to maintain them. A response pattern of this kind is involuntary and respondently conditioned. Many emotional responses are initially of this type. Most responses with which educators show concern, however, are not involuntary but rather are conditioned through a history of reinforcement. That is, their patterns of responding have been strengthened by stimulus events in the environment functioning as consequences of behavior. Skinner has coined these operant responses, for as they occur they operate on the environment. When an operant response occurs, it has the effect of making a change in the environment and that change acts as a consequence for the response which operates on it. This consequence, when it functions as a reinforcer to the individual, serves to strengthen the responses it reinforced, the type of response with which the remainder of the behavior principles are concerned.

Through specific arrangements of consequences following a specified response, the rate of occurrence of these responses can be increased, decreased, maintained, or extinguished predictably. The principle of positive reinforcement explains the effect of a pleasant event in strengthening the probability of the occurrence of the response it follows. If the teacher's attention, or her smile, or her statement of "good job" is a pleasant event for a child, the teacher can react to the child in one of these ways following a pattern of academic performance and predictably accelerate the child's rate of performance. The principle of positive reinforcement can be viewed as a very general principle incorporating a number of sub-principles, all describing either (a) types of consequences which function as general positive reinforcers, or (b) schedules for presenting reinforcement which bring about precise patterns of behavior. It is not enough simply to present a pleasant event sometime after a pattern of behavior or a set of responses has occurred, in order to establish the behavior efficiently. Acquisition of a response occurs most predictably when reinforcement is immediate and continuous. When a high rate of the behavior has become established, then reinforcement need occur only intermittently.

Specific types of stimuli come to acquire strength as positive reinforcers and can be described in terms of the strength and generality they predictably acquire to influence behavior. The principles of conditioned reinforcement and generalized reinforcement explain the environmental arrangements which establish a wide variety of objects, events, conditions, and our own responses as pleasant events which can be used to increase the probability of responding. The human smile, the pat on the back, the words in a book, are not initially events which strengthen behavior, although for most individuals these stimuli gain strength when paired systematically with consequences already pleasant.

Negative reinforcement is a principle of behavior describing conditions which strengthen the probability of the occurrence of a pattern of responses through removal of an aversive stimulus after a response--arrangements which lawfully produce escape and avoidance behaviors. Children in the classroom who never begin working until the teacher becomes very stern and scolds or nags have behaviors controlled through negative reinforcement. These children typically stop work soon after the teacher stops prodding.

The principle of extinction describes environmental conditions that predictably eliminate a pattern of behavior. Arranging events so that a positively reinforcing consequence no longer follows a particular response pattern is the operation that leads to the elimination of that behavior. For example, if the teacher will cease to attend to the child when he is shouting out or leaving his seat unnecessarily, these behaviors will decrease in number and eventually disappear, if it is the teacher's behavior that is maintaining it. Extinction occurs most effectively when a response incompatible with the response being extinguished is concurrently reinforced.

Scheduling the occurrence of reinforcement is as important as the type of reinforcement presented. Behavior is concurrently influenced by both. Attention to the scheduling of reinforcement is critical to the acquisition of a new pattern of behavior as well as to the maintenance of a strong and stable behavior pattern over long periods of time.

For most children in almost all classrooms, acceptable rates of academic responding have been shaped and are maintained by natural reinforcing events from the classroom. There are other children, however, who exhibit a different reinforcement history, children who are described as lazy or apathetic or who dislike reading or hate school. As with the children whose behaviors appear very acceptable, these behaviors could be explained if a record of the child's reinforcement history were retrievable.

Procedures of Experimental Analysis

Experimental education, because it begins from the foundation already laid by the principles of instruction and behavior, should not imply "new and untried methods." Rather, education in the form described here is experimental because of the degree of control over classroom conditions that is possible with the use of procedures of experimental analysis. Common use of procedures of experimental analysis will facilitate the growth of experimental education. Through the use of experimental analysis, the newly obtained information can then be used to apply to information already reliably obtained with similar procedures, thus providing systematic replication of the original findings and, consequently, systematic replication of the new findings.

Classroom Application

The procedures of experimental education are synonymous with the procedures of experimental analysis as they are used to apply and extend the principles of instruction and behavior. Specifically, these procedures are characterized by direct observation of the dependent variable (behavior), continuous measurement of its occurrence under the controlled conditions established, and systematic manipulation of the independent variables to be investigated. The dependent variable--always a response well defined by its observable topography--is measured by its rate of occurrence. Rate is the basic datum and its record provides a sensitive tool for predicting the probability that a specific behavior will occur under certain conditions. The independent variables--always stimuli well specified by their observable dimensions--are systematically investigated to determine their influence on the probability of the occurrence of the specific behavior. The objective is to study behavioral processes as they are observed in changes in rate of responding (behavior patterns) due to the function of variables systematically manipulated.

Direct observation is systematic observation of behavior, usually involving several degrees of refinements in observation procedures which may first begin with a narrative description of the behaviors observed and then identification of specific behaviors to measure. Once the behavior selected for further observation is identified it is defined by its precise topographical unit or cycle to permit a tally of its frequency. The occurrences of these units of behavior are then counted over time in order to determine the rate of occurrence.

Continuous measurement of these responses requires that a response topography be selected which will maintain its comparability during changes in environmental conditions even though response requirements increase in difficulty as academic materials naturally become more complex. This facilitates the sensitivity of measurement necessary for precise evaluation of the effects of changes in contingencies and reinforcers.

Systematic changes in environmental conditions enable the evaluation of event changes which lead to the establishment of prescribed behavior patterns or sets of responses. Because behavior is lawful, and because it develops lawfully from environmentally arranged conditions, the influence of these conditions can be determined if changes are introduced one at a time and held constant while measurement of performance is taken. A pattern of behavior may not initially register the effect of the temporally arranged conditions. Therefore, that condition must remain as introduced over a period of time for reliable evaluation of its influence.

Applying principles of instruction and behavior in the classroom, using procedures of experimental analysis, is commonly referred to as contingency management. A relatively recent innovation in the classroom, it is defined most precisely in terms of the systematic utilization of reinforcing events in relation to specified behavior. However, three classroom variables are relevant to the contingencies responsible for changing behavior: (a) the occasion upon which the behavior occurs, (b) the performance of concern, and (c) the consequence of behavior. Armed with this important information, the teacher can have a strong and predictable influence upon behavior by arranging conditions which facilitate the establishment of appropriate classroom behavior. Experimental education, as it influences instruction in the classroom through research, service, and professional training, is directed toward this end.

Four Components of Instruction

For the teacher, experimental education is characterized by the responsibility for four components of instruction. To conduct effectively her responsibility for the academic progress of each student, she must attend to cueing, response measurement, reinforcement, and contingency management.

Cueing. The first major responsibility familiar to every teacher and taught within every college of education is the task of presenting material to the child to bring out the kind of responses he must make to develop a specific skill. Cueing is the basis of the modern curriculum and has been the focus of educators historically. Cues are of many types, depending upon the task requirement, are presented in many forms, and are received through the senses auditorically, visually, or kinesthetically. The critical feature of cueing is the sequential arrangement designed to increase the probability of accurate responding relevant to each skill level.

Response measurement. The teacher not only provides the cues, she must also measure the responses the child makes to these cues. Measurement procedures range from a simple count made by the teacher or child to very complex recording of the responses and temporally occurring events.

Reinforcement. The third phase of the teacher's responsibilities is reinforcement, a procedure which involves planning and presenting specific events to follow a type of responding. Educators are rapidly becoming aware of the many other conditions in the classroom which influence skill development. Conditions or events which follow a response have a direct influence on behavior whether planned by the teacher or not. They may be aversive or pleasant and include anything relevant to the child in question. To each child, several kinds of events may serve as reinforcing consequences for his behavior: teacher attention, peer attention, a new assignment, free activity time, or any number of other kinds of work or play activities common to the regular classroom. Furthermore, a consequence for responding, presented to the group, will be positively reinforcing for some children, neutral to some, and aversive to a few. Systematic attention to consequating the child's responses, therefore, is as important a teacher responsibility in skill development as is systematic attention to cueing.

Contingency management. The contingencies the teacher establishes between the child's performance and events which follow immediately are the procedures which modify behavior--the fourth component of instruction. For effective contingency management, all the classroom conditions impinging on the child's performance must be identified and held constant while the child's pattern of responding, under the prevailing contingencies, is measured over several sessions and then compared to the child's baseline performance patterns.

With these procedures, the relationships existing between the behavior and environmental conditions can be described in terms of rate of response. In addition, the parameters of any dependent variable, as they function in relation to the behavior, can be described in terms of changes in rate of response.

Through the realization of experimental education in the classroom, teachers will plan and conduct instructional programs where they systematically arrange and present classroom events in temporal relationships with child behavior to facilitate performance. The teacher will also incorporate into daily activities the scientific procedures for measuring the performance of the pupil in order to evaluate the effectiveness of instructional arrangements.

Conclusion

Experimental education, therefore, is characterized by three inherent features: (a) a focus on behavior, (b) instructional procedures based on the principles of instruction and behavior using procedures of experimental analysis, and (c) instructional decisions based on the response data of the child obtained under conditions where the teacher had some degree of experimental control. The substantial and rapid accumulation of experimental evidence supporting the lawful relationship between the observable unit of behavior and the temporally related events has become a significant point of view on behavior acquisition. Broadly encompassed within these dimensions, education now has a framework from which it can develop as a major scientific discipline in the study of classroom behavior. Out of this framework for a major scientific discipline, the clinical practice of education can become more precise and efficient.

References

Skinner, B. F. Science and Human Behavior. New York: MacMillan, 1953.

Wallen, N. E. and Travers, R. M. W. Analysis and investigation of teaching methods. In N. L. Gage (Ed.) Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Warren, A. B. and Brown, R. H. Conditioned operant response phenomenon in children. J. Gen. Psychology, 1943, 38, 181-207.

CHAPTER III ASSESSING THE EFFECTIVENESS OF INSTRUCTIONAL PROGRAMS[1]

A. A. Lumsdaine

This chapter is necessarily addressed to several audiences. Chief among these are (a) the program user or potential user interested in determining the suitability of a given program for his educational purposes; (b) the program producer, interested in providing data to attest to the merits of the programs he hopes to market or otherwise distribute for use; and (c) the behavioral scientist or educational technologist who, in addition to other interests, may be able to provide technical assistance to the user or producer in obtaining or interpreting assessment data.

Since the background and interests of these three groups may differ considerably, some compromise is necessary if the chapter is to be useful to all three. The attempt is made here to discuss major issues in a sufficiently simple, nontechnical manner to be intelligible to the seriously interested nontechnical person concerned with program assessment, either as user or producer, while also trying to identify some of the more important technical problems involved.

Reprinted from Teaching Machines and Programed Learning II: Data and Directions, edited by Robert Glaser, published by the Department of Audiovisual Instruction, National Education Association, 1201 Sixteenth Street, N.W., Washington, D.C. 20036. $11.50 per copy. Stock No. 071-02390.

Distributed by the Joint Committee on Programmed Instruction of the American Educational Research Association, American Psychological Association, and Department of Audiovisual Instruction, National Education Association, with the cooperation of the Educational Media Branch, Office of Education, U.S. Department of Health, Education, and Welfare under the auspices of Title VII, Part B, of the National Defense Education Act.

The Problem of Assessing Program Quality

The problem with which this chapter is concerned was anticipated in the following remarks, written in the spring of 1960: "In the production of programs a major problem could arise from premature publication and sale of hastily conceived and untested programs. . . . It would therefore appear that a high-priority objective is that of working out acceptable quality-control standards for programs" (Lumsdaine and Glaser, 1960).

Concern with evaluative criteria for assessing the quality of programed materials was primarily responsible for the formation, in 1961, of the Joint Committee on Programed Instruction (J.C.P.I.), representing the American Educational Research Association (AERA), the American Psychological Association (APA), and the Department of Audiovisual Instruction (DAVI) of the National Education Association (NEA).[2]

Many other individuals and groups have also been concerned with this problem. In addition to the J.C.P.I. reports (AERA, 1961, 1963, 1964; Lumsdaine, 1962c, 1963c; Ryans, 1961; NEA, 1964) and previous papers by the present author (Lumsdaine, 1962a, 1962b, 1962d, 1963a, 1963d), discussions of the problem of program assessment have been provided by Geis (1962), Eigen (1961, 1964), Gotkin (1963), Rothkopf (1961, 1963), Silverman (1964), Stolurow (1964), Caulfield (1963), Schutz, Baker, and Gerlach (1964), Holland (1961), Glaser (1963), Hively (1964), Maier, Stolurow, and Jacobs (1963), and others. The paper by Lumsdaine (1963d) presents a more extended discussion of some of the methodological problems encountered in assessing and describing the effects of program use.

Background and Perspectives

In 1961, the AERA Joint Committee pointed out that the contribution of self-instructional programed learning materials, used in teaching machines or otherwise, can be best realized only if users have adequate information with which to evaluate programed materials. Some of the interim guidelines prepared by the Joint Committee in 1961 are relevant as perspective for the present discussion. The concluding statement is as follows: "The effectiveness of a self-instructional program can be assessed by finding out what students actually learn and remember from the program. The prospective purchaser should find out whether such data are available and for what kinds of students and under what conditions the data were obtained" (AERA, 1961). This statement suggests the perspective reflected in the main concern of the subsequent work of the Committee--namely, the assessment of individual instructional programs in terms of their demonstrable performance characteristics. In its second published report (AERA, 1963), the J.C.P.I. further developed this perspective and amplified the foregoing recommendations. The points of view given in this report, quoted several times herein, also represent a basic perspective for the present paper.

Product testing vs. evaluation of a "method": A crucial distinction needs to be made between the question of assessing the quality of specific programs and the question of evaluating programed instruction as a general method. This chapter is exclusively concerned with the former question, considered as a useful form of *product* assessment. However, the restriction of product-assessment studies to the immediate aim of determining the quality or suitability of a particular program (with no attempt, as a primary objective, to derive generalizations about the methods represented) does not preclude the possibility that leads about such generalizations may emerge as important by-products of these studies (Hovland, 1949). Evaluation of programed instruction as a general method is a much more difficult and elusive question to answer. This is so because of the difficulty of defining the "method" of programed instruction in general terms, or of delimiting it--as well as alternative "methods"--in a way that would provide a basis for a generalizable answer for a question stated in such nonspecific terms.

The need to distinguish between assessment of a particular program and of the "method" it purports to represent has been stated in the 1962-63 AERA Joint Committee report (1963):

> . . . the value of a method of instruction cannot be tested in the abstract. For example, evaluation of a particular textbook is not an assessment of the usefulness of textbooks in general. A properly constructed experimental tryout or field test of a program may provide an assessment of that particular program, but does not afford proof or disproof of the value of a general "method" of programed instruction.
>
> Experimentation conducted thus far supports the expectation that good programs, carefully developed, *can* significantly

> improve the quality and economy of instruction. Whether any particular program *will* do so is subject to question until established by adequate tests of that program

Merely recognizing this point does not, of course, insure its comprehension by the program-buying public. One of the problems which can only be solved as data on performance characteristics for each specific program are made widely available is the "halo" which boils down to the following invalid syllogism (examples of which, in hardly less blatant form, have appeared widely in advertising copy):

> Some programs have been shown to teach very effectively;
> These materials which I offer you are programs;
> Therefore, these materials provide a superior way to teach your students.

In a continued attempt to combat the tendency to accept such spurious arguments, the J.C.P.I. has again highlighted in its 1964 report the need to judge each program individually, by restating the point as its first recommendation to prospective users: "Prospective users should evaluate each program on its own merits according to its demonstrated effectiveness rather than relying on general statements or findings purporting to support the value of the 'method' of programmed instruction" (AERA, 1964). The attempt to assess the general worth of any "method" or "medium," including programed instruction, really involves an essentially meaningless question. As has been elaborated elsewhere in more detail (Lumsdaine, 1963b), attempts to compare any medium or method with another in the abstract, so as to support a generalization about the value of the medium or method, are inherently foredoomed to failure for the simple reason that a good film, for example, will always beat a poor lecture, and vice versa. Meaningful experiments thus must either have the purpose of determining the effects produced by specific programs or must seek to test propositions about the effects of definable, describable properties of programs.

The difference between this latter purpose and that of assessing specific programs reflects the distinction between the scientific and the technological goals of research and development on instruction. The technological goal is concerned with the development and description of demonstrably good *products*; the scientific goal comprises the generating and testing of hypotheses which can lead to the development of *principles*, ultimately comprising a science of instruction (Lumsdaine, 1961, 1962e, 1963b, 1964). It is to be emphasized that the important long-range contributions in programed instruction will result from scientifically oriented studies which seek to identify and validate rules or principles of programing that transcend the properties of specific programs. The important short-range efforts on which this chapter focuses are, by contrast, directed at ascertaining the quality of specific individual programs in terms of what their use can contribute to specified instructional outcomes. Scientifically oriented studies for testing hypotheses or proposed principles of programing can be considered here only incidentally, insofar as they affect the choice of methods used in the assessment of specific programs.

Importance of Program Assessment

The State of the Art in Program Production

The author is convinced that most existing programs afford only a rough approximation of the potentiality for control over learning which could, in principle, relize a goal of assured mastery for all qualified students. This position can be argued both on a priori grounds and in terms of such limited data as are currently available on the effectiveness of existing programs (Bolt, et al., 1963; Drooyan and Wooton, 1964; Fletcher, 1964; Glaser, 1963; Glaser, et al., 1963; Paulson, 1963; Schramm, 1964a). Even casual inspection of a sample of programs suggests a tendency merely to follow superficially the general format implied by one programing rationale or another, while meeting neither the theoretical assumptions nor empirical characteristics that are supposed to be exemplified. In addition to lack of adequate tryout and revision, many other apparent weaknesses are to be seen in examining the existing programs, including inadequate analysis of subject matter content and inept use of what seem to be the more promising techniques of programing. Accordingly, it should not be surprising if, despite the acclaim accorded to programed instruction as a basis for a potential "educational revolution," many current programs do not prove to be more effective than alternative kinds of instruction. The existence of a gap between the promise of programed instruction and its realization, up to 1962 at least, is a major thesis of the provocative report by Schramm (1962), who defends the position that while "programmed instruction is, in the best sense of the word, a truly revolutionary device," its "potential is, so far, largely unrealized." (See also Rothkopf, 1964).

Knowledge of what programs are available as a basis for choice: An elementary step in assessing any program, particularly in terms of its content, is simply to know what other programs are available in the same or similar subject matter. The publication of the USOE-sponsored survey edited by Hanson at the Center for Programed Instruction (1963) and the compilation by Hendershot at Delta College (1963) have been helpful in this respect. But mere knowledge of the availability of programs, while at least showing the prospective purchaser that he may need a basis for choosing among available alternatives, does not provide him with standards of judgment for making the choice.

Why "Standards" for Assessment?

The question may well be asked: Why have "standards" or "criteria" for assessing the quality of programs? Why are such criteria desirable, feasible, or justified, as compared with the case for other instructional resources like textbooks, films, simulators, or other training devices? Attempts have long been made to develop criteria for evaluating films, training devices, and other instructional tools (Edgerton, 1960; Los Angeles County Board of Education, 1963; Maier, et al., 1963; Miller, 1953; Stolurow and Lumsdaine, 1956). The main differences between these previous attempts and the problem as considered here lies in the attempt to develop validating criteria based on controlled measurement of what the use of a program demonstrably contributes to the attainment of behaviorally specified instructional goals. As the AERA Joint Committee has pointed out:

> The tendency to empirically guided development of programs is coupled with an orientation toward testing the specific effects produced by a program, and toward more sharply focused objectives defined in terms of specified behavioral outcomes. In addition, the program is intended to generate a more predictable pattern of student behavior than does the study of a textbook, which generally has a less specialized purpose in aiming to serve as a reference source as well as a sequence of instruction (AERA, 1963).

The usefulness of criteria for assessing program effectiveness: A basic purpose for developing criteria to assess the quality of specific programs is to increase the usable potential of programed instruction, both through improving the selection and use of existing programs and through stimulating the development of more effective programs in the future. Both the wise selection and the effective utilization of present programs in schools clearly requires a dependable way to assess the merit of any given program.

Effect of standards on program production: Part of the case for introducing dependable and widely accepted criteria of assessment lies in the effect on the standards of quality in future programs, particularly those produced by commercial publishers. We may assume that a stimulus to better quality production in this field, as in others, involves the dynamics of a competitive marketplace. If the consumers (e.g., school systems) have a dependable method for differentiating better programs from poorer programs, a demand for the former is effectively generated, and publishers must produce better programs in order to compete in the market. However, such competition cannot be effective unless there is indeed a basis for determining the quality or effectiveness of programs in unambiguous terms. In the absence of available unbiased and dependable information about program quality, programs can be promoted and sold on the basis of unsupported claims or dubious "data" purporting to show their merits, and the competitive incentive to produce genuinely superior programs is thereby weakened. Stimulation of program quality by the open competition of the marketplace is next to impossible in the absence of dependable and acceptable criteria for assessing the merits of any particular program (Lumsdaine, 1963a).

What Kind of "Standards" are Relevant?

Some Basic Distinctions Among the Main Kinds of Criteria

Three main kinds of considerations need to be distinguished as relevant bases for assessing the suitability or acceptability of a particular program for meeting a given educational purpose: These may be termed "appropriateness," "effectiveness," and "practicality."

"Appropriateness," as used herein, refers to the nature of the "subject matter" or "content" that is "covered" by a program. The concept of "content" actually turns out to be a rather fuzzy and unsatisfactory one, with some ambiguous and troublesome connotations. For the present purpose, program content can be characterized as representing what the program tries to teach or, perhaps, what it "contains" that apparently could be learned by an optimally qualified student who learned everything that it was possible to learn from what is presented in that particular program. In other words, appropriateness may refer to prospective outcomes to which a program's use might lead, that is, to what is to be learned or may be learned from a program. Thus, appropriateness means, roughly, the extent to which program "content" is consonant with the objectives of a particular educational purpose or course, or the degree of correspondence between the user's objectives and those of the programer.

"Effectiveness" refers to how well the program does, in fact, attain certain prospective outcomes, how well it teaches whatever it is calculated to teach (rather than *what* it may teach), or, in other words, the extent to which its content is learned or the extent to which stated objectives are attained by students who use the program in a particular way. A further distinction can also be made between effectiveness and efficiency, the latter referring, broadly, to the extent to which a given degree of attainment is achieved economically in terms of the use of student time and other resources.

"Practicality" can be used to refer to matters of cost, feasibility, acceptance by students and teachers, and other factors which determine whether an appropriate program of given potential effectiveness can or will in fact be used so that its potential is realized in practice. This category involves, aside from factors of convenience that may influence effectiveness, considerations that are largely translatable into terms of cost.

Finally, the terms "suitability" or "acceptability" might be used in a generic sense to indicate over-all bases for evaluation or decision concerning program adoption or use, based on consideration of all three of the above classes of factors (appropriateness, effectiveness, and practicality).

Interrelation Between Appropriateness and Effectiveness

Obviously, both appropriateness and effectiveness are important considerations in assessing a program. Almost as obviously, they do not necessarily go hand in hand. As pointed out by Galanter (1961), a program might be effective in teaching inappropriate content, or it might present appropriate content but fail to do so effectively; also, it might teach inappropriate content either effectively or ineffectively. One reason for largely bypassing the question of appropriateness as a primary concern here is simply that the determination of appropriateness is such a complex problem and involves many unsettled questions of value in terms of what should be taught. At present, at least, it can be held that each user or reviewer can claim to be as good an authority as the next.

However, even though primary emphasis is placed on the effectiveness with which a program teaches, the question of what the program is supposed to teach, and hence what should be measured in determining its effectiveness, will necessarily enter into this discussion to some extent. Furthermore, it should be recognized explicitly that "assessment" of program effects, in the sense of their measurement and description, clearly does not in itself provide "evaluation" of a program; at best it only provides an important basis on which, along with other relevant information, an evaluation can be made of the suitability of a program for meeting a given set of instructional objectives.

"Internal" and "External" Sources of Information About Programs

A useful terminological distinction suggested by Silverman (1964) and by Rothkopf (1961), which was also adopted by the AERA-APA-DAVI Joint Committee (AERA, 1963), can be made in terms of the locus or source of information about a program. This is the distinction between internal and external sources of information as possible criteria for program evaluation. "Internal" characteristics refer to features which can be revealed through inspection of the program material, including both its "content" and such pedagogical features of construction as length of frames, use of branching, techniques of prompting, patterns of repetition and review, kinds of responses called for, and the like. These may be viewed merely in a descriptive sense, but often are assumed to be predictive of the effectiveness of the program. Clearly, if there were a fully developed science of instruction, the effectiveness of a program could be predicted by determination of the extent to which such descriptive characteristics of a program were optimally selected and arranged to promote effective learning.

"External" information about a program refers to features which cannot be observed merely by inspecting the program itself, such as the history of the way in which it was developed and, in particular, its observed performance as a teaching instrument. Other kinds of external information could include such information as the qualifications of the author, the kind of student-response data obtained in revising the program, opinions of reviewers, and test data obtained to measure the achievement produced by the program.

Predictive vs. Validating Criteria of Effectiveness

A further important distinction can be made between (a) those external criteria that are believed to be *predictive* of program effectiveness (such as external evidence about the competence of the programer or the history of a program's development, including tryout and revision) and (b) *validating* criteria, which consist of direct evidence of the effects actually produced by the program in demonstrably changing students' behavior.

Experienced programers will undoubtedly continue to look at programs and state, possibly with some real basis for confidence, that they are good programs or poor ones. But this is an unvalidated opinion, though it may be an informed and illuminating one. It is a prediction of effectiveness, not a verification. It should not be greatly surprising, therefore, to find that some programs that looked poor may turn out to do a good job of teaching, or that frames which seemed beautifully fashioned may, when put to the test, do a poor job of teaching. The distinctions among major classes of criteria for judging a program's effectiveness (or efficiency) may thus be reoriented as follows:

1. Predictive criteria of effectiveness: rational or theoretical bases, involving inferences from general experience or extrapolations from laboratory science, on which the effects of a program are believed to be at least partly predictable. These include (a) *internal criteria*, derivable from inspection of the program, and (b) *external predictive criteria*, based on ancillary informa- about a program's development, or on external information such as expert review or information about students' or teachers' opinions of a program.

2. Validating criteria of effectiveness: external criteria that provide measures of the actual effects of the program in demonstrably changing students' behavior.

Possible "Validating" Criteria for Assessing Program Effectiveness

Validating criteria have been characterized as measures of the actual effects of the program in changing students' behavior. The validity of such measures may vary, however, along a dimension of ultimate validity from responses the student makes within the program to measures of ultimate or long-term retention, transfer, or application. Measurement of such effects may also vary with respect to inclusiveness of *all* relevant effects (including transfer and motivational and other "indirect" effects as well as direct competence per se on the subject matter covered) and also with respect to how clearly the behaviors observed represent changes demonstrably shown by rigorous experiment to result from the use of the program. For example, the following kinds of evidence differ from each other in one or more of these respects.[3]

1. Error rate on prompted frames or over-all error rate.
2. Error rate on frames that are internally unprompted, but are located within or just following prompted program sequences, so that "sequence prompting" effects are present.
3. Error rate, or pattern of errors and correct responses, on review sequences placed so that they are minimally effected by sequence prompts.
4. Gains from preprogram to immediate postprogram tests.
5. Gains from preprogram to immediate postprogram tests, but with control for external influences.
6. Demonstrated changes on indirect motivation and transfer.
7. Persistent or "permanent" effects as shown by delayed tests of transfer and application, including sustained motivation.

The major factors that bear on deciding the suitability of a program are summarized, in relation to sources of information concerning them, in Figure 1.

Critical Reviews of Program

In recognizing various levels of assessment for program quality, it is necessary to look further at the possibilities of critical reviews of pro-

	Internal Information (Available from program inspection)	*External Information*
Appropriateness:	Content inspection	Stated objectives
	Table of contents	Test content
	Reading of Program	Competence of authors
	Analysis of terminal-behavior frames	Opinions of reviews or advisers
	Publisher's statements of objectives and tests provided by puboishers	
Effectiveness:		
1. Predictive criteria	Features of program style or construction	Reviewer's opinions
	Inferred direct and side effects based on program inspection	Developmental history, including tryout and revision
		Error rates within the program
2. Validating criteria	----------------------	Measured effects of program use and related data (time, etc.)
Practicality:		
	Ease of using	Cost factors: program prince, adaptability, characteristics of presentation, machine (if required)
	Reusability	
	Machine (instrumentation) requirements	

FIGURE 1. Main Types of Criteria for Assessing the Suitability of a Program.

grams, which furnish one possible basis for evaluation. Such reviews began to appear in professional journals starting around 1961. Examples are reviews by Eigen (1961), Galanter (1961), Saltzman (1961), and Silberman (1961) on algebra programs, by Markle (1961) on a program in English grammar, by Carroll (1964) on a program for teaching Russian script, and by Denova (1963) on a program for teaching digital computer programing. Other periodicals in which reviews of programs may be found include Audiovisual Instruction, AV Communication Review, and the NSPI Journal (National Society for Programmed Instruction). As with reviews of tests, some program reviews include data--in this case, data on achievement attained by using the program--as well as reviewer opinion about the program based on its internal features (Carroll, 1964; Denova, 1963; Fletcher, 1964; Galanter, 1961).

The emphasis on objective standards for assessing programs, on which this chapter is focused, should not minimize the potential usefulness of critical review based on inspection of programs. However, aside from the need to assess the competence and bias of reviewers, users should be made aware that reviewer opinions may conflict, and furthermore that no reviewer may correctly predict what the program will actually teach. The usefulness of reviews will thus be increased as provision is made for the collating and bringing together of several reviews on each program so as to have available something like the collection of reviews of tests provided by the O. K. Buros' Mental Measurements Yearbook (Buros, 1959a; Buros, 1959b).

Program reviews, even when only representing reviewer opinion, can be helpful in view of the need to make recommendations and decisions about the acceptability of programs in the absence of objective data about what a program's use can accomplish; lacking such data, one has to depend primarily on opinions of reviewers based on program inspection. Though opinions clearly do not qualify as criteria in terms of which programs can be objectively assessed, they may be viewed as signposts that are useful if accepted as advice rather than fact; they offer something to rely on "until the data comes," particularly in assessing the appropriateness of program content. Whatever its value, it seems certain that as with other educational materials (e.g., textbooks, films), reviewer opinion will be used as one basis for evaluation of programed materials.

As objective and valid data on demonstrated program effectiveness becomes increasingly available, the main function of reviews may be to provide a critical analysis of the validity of results from objective assessment studies and to furnish advice on over-all suitability of programs, taking into account content appropriateness and practicality of use as well as demonstrated effectiveness indicated by experimental data. Reviews, even by programing "experts," necessarily represent predictions of program effectiveness rather than objective evidence of it, except when based on data from objective studies of program effects. Data offered by a reviewer in support of conclusions about a program's effectiveness should, moreover, be weighed in the light of technical considerations influencing the validity of such data, discussed later in this chapter. Data for two or three students gathered informally by the reviewer may be indicative, especially for extremely poor or extremely good programs, but do not take the place of more formal and extensive assessment studies.

Guidelines for reviewers: The J.C.P.I. (AERA, 1964) has recommended that those who prepare critical reviews of programs should, in addition to expressing their opinions about the suitability of the program content and objectives: (a) obtain and report all available data about program effects; (b) evaluate and interpret such data in the context of technical considerations such as those set forth by the Joint Committee; and (c) distinguish clearly and explicitly between their own opinions about the probably effectiveness of the program and the objective evidence on its demonstrable outcomes.

Checklists and Other Statements of Proposed Evaluative Criteria

Many checklists and statements of criteria for assessing programs have been proposed by a number of sources.[4] These have tended to represent a potpourri of criteria related to appropriateness, practicality, and both internal and external predictive criteria of effectiveness, together with external validating criteria (i.e., measured program effects). In such statements and checklists, little explicit differentiation or recognition has been made of the status of the differences in kinds of criteria proposed in terms of the foregoing kind of distinctions. However, the distinctions seem to be useful ones even though the three primary classes of factors may interact and in some ways overlap. As one example of this overlap, in the above-noted distinction between effecriveness and efficiency it is evident that the latter involves cost factors, indirectly at least; also, on closer examination, the meaning of "content," as commonly used, will be found to be related both to appropriateness and effectiveness.

In the course of preparing this chapter, the writer and some of his students examined and attempted to classify several hundred statements put forth in published documents as criteria for the assessment of the suitability of a program in terms of its appropriateness and/or its presumed effectiveness. Most of these statements came from "checklist" formulations put forth for the guidance of parents, teachers, curriculum specialists, etc. (Beltron, 1962; Center for Programed Instruction, 1962; General Programmed Teaching Corporation, 1963; Jacobs, 1964; NSPI, 1962; NSPI, 1963; New York City Board of Education, 1962; Rocky Mountain School Study Council, 1962; Teaching Materials Corporation, 1962; Tracey, 1963; University of Michigan, Center for Programed Learning for Business, 1963; U.S. Air Force, 1962). These statements can be grouped in several broad categories. A considerable number of them refer to internal characteristics of the programs, either to factors of construction and organization presumed to be predictive of effectiveness or statements in which the above-noted overlap between effectiveness and appropriateness makes an unambiguous classification in this respect difficult or impossible. A second group of statements refers more unequivocally to appropriateness factors, either in terms of what is to be taught or the kinds of students for whom the program is appropriate. Another group of statements concerns questions of feasibility in pattern of use, questions which may, depending on point of view, be considered to relate either to the appropriateness or to over-all effectiveness in school use in a variety of use patterns. A smaller group of statements refers to external characteristics, particularly the history of developmental testing, tryout, and revision which the program has undergone. A final category refers to external validating data and their interpretation. There is often some ambiguity between developmental and descriptive or validation data due to vagueness in the way the data are reported; sometimes it is not possible, for example, to know whether "tryout" data refer to information used as a basis for revising the program or presented to attest its effectiveness.

Some such checklists seem to imply the possibility of deriving a "score" for a program, in which acceptability for a program can be determined from the number of favorable answers to the questions posed. In the opinion of the writer, such an implication is a mischievous one, particularly since there is no assurance of the validity of many of the questions asked (particularly those about internal characteristics). Even where the questions are clearly relevant, there is no assurance as to how they should be weighted. It would be interesting to apply such questions in systematic fashion to the effects actually produced by a number of parallel programs with similar objectives; this might serve to determine whether, regardless of theoretical rationale, they appear to have any empirical predictive validity (cf., Rothkopf, 1963).

It is also interesting, with respect to questions which appear to be clearly relevant, to consider the order in which it is most appropriate to ask questions. Such ordering has been implied to several checklists, including that of ETS (Jacobs, et al., 1964). One might devise a kind of decision flowchart, algorithm, or structured "20-questions" game, on the basis of which one could examine programs efficiently. Such a 20-questions arrangement, in the figurative sense, would differ from the "20 questions" proposed by Belton (1962), which do not form an algorithm, but merely a checklist of points to be considered. Undoubtedly some kind of spiral or alternation between several major categories of consideration--appropriateness, feasibility (including cost), and probable or demonstrated effectiveness--would be reasonable in considering the adoption of a program. One might ask a few over-all screening questions: for example, whether content appears to be at least "in the right ballpark" and whether its cost is conceivably feasible, etc., before proceeding to more detailed examination of the program in terms of presumed or demonstrated effectiveness and more detailed aspects of content suitability. If neither of these questions could be answered in the affirmative, the program would be ruled out for further consideration.

Any criteria for determining the suitability of any course of action, including the adoption of an instructional program, involves matching available means to desired ends. It follows that any such criteria must include a specification of the ends sought by the user. Neither the specialist in instructional programing nor the publisher of programs has, as such, any special competence, much less authority, to tell the user what his aims should be. The user must decide these for himself. In doing so, however, he may perhaps wish to examine the objectives which the programer has formulated in writing the program or other statements of possible outcomes relevant to the general field to which the program pertains.

Primary Reliance on External Validating Criteria

The notion of "standards" of effectiveness has suggested to some the development of authoritarian or restrictive criteria which attempt to dictate the way programs are written or presented. This unintended and quite unfortunate connotation has tripped off various tirades against the attempt to develop criteria (e.g., Esbensen, 1962). Any such attempt to standardize or freeze program styles would be very undesirable, as was stressed in 1960 by Lumsdaine and Glaser:

> In the development of quality-control standards for programs, it is important to avoid the imposition of inflexible requirements which might inhibit creativity and experimental use of new techniques.
>
> It seems clear that standards for the adequacy of a program ought to be conceived primarily in terms of its effectiveness in attaining defined educational objectives, rather than by specifying the format, sequencing, or other aspects of the means whereby these ends are achieved (Lumsdaine and Glaser, 1960, p. 566).

This emphasis on avoiding any prescription of internal form or style and advocacy of external, validating criteria as the primary basis for assessing the effectiveness of programs has also been consistently advocated by Rothkopf (1961) and the AERA Joint Committee (AERA, 1963, pp. 87-89; see also Silverman, 1964). It is the empirically oriented position that ultimately the "proof of the pudding is in the eating," that is, that the ultimate measure of a program's effectiveness is what it teaches.

<u>Dangers of restrictive "standards" based on internal criteria which would prejudge program effectiveness</u>: Schramm (1962) has pointed out the existence of a tendency toward premature "freezing" of particular programing styles. This fixing on a stereotyped style can be seen in many current programs, despite the warning given five years ago in the statement

just quoted. Undoubtedly, it reflects an overpreoccupation with internal criteria inferred from early programs and a tendency to imitate their superficial characteristics rather than experimenting with new styles and relying on empirical proof to determine how well the resulting programs work.

Rationale for reliance on validating criteria: The decision to limit criteria of effectiveness in this chapter's discussion of program assessment to validating criteria or measured effects produced by programs is suggested and made possible by the conception of programs as potentially autonomous vehicles of instruction. Such forms of assessment have not characteristically been applied to textbooks or other instructional materials. It is the tendency for development of programs to be based on an explicit statement of objectives--and for programers to take the responsibility for achieving these goals without dependence on other forms of instruction--that makes possible a policy of accepting empirical data as the validation of the program's effectiveness. There is some similarity here to the rationale underlying empirical validation of psychological and educational tests, as has been pointed out by the Joint Committee (AERA, 1963). Although programs and tests differ in objectives, with programs aiming primarily to instruct rather than to test students, both generate student-response data and are capable of being developed as well as validated in terms of empirical procedures. In both cases an external criterion can be specified, at least in principle, to indicate the extent of which an intended outcome has been achieved as evidenced by kinds of behavior which have been developed (in the case of a program) or differentiated (in the case of a test).

The risk in relying on inspection for assessing program effectiveness is that widely accepted precepts and current patterns of programs have not, as yet, been the subject of satisfactory experimental validation. Although some considerable number of experiments comparing the relative effectiveness or efficiency of alternative forms of programs have been conducted, inspection of the available evidence makes it clear that a great deal more evidence than is now available is necessary before a well-developed science or validated theory of programing, on the basis of which program effectiveness can be reliably predicted, can be delineated (AERA, 1963; Rothkopf, 1961; Rothkopf, 1963).

One form of evidence bearing on theories or principles of programing comes from comparative studies in which controlled variation of specific program features has been introduced. For example, the importance of having the student overtly compose responses, as stressed by Skinner (1958) and others, has been studied in a number of investigations (see Lumsdaine, 1961; Lumsdaine and May, 1965; Schramm, 1962; Schramm, 1964b). In general, such studies have come rather far from offering clear-cut support for the principles of programing which suggested the alternative forms of programs that were experimentally contrasted. Although many of the experiments thus far performed suffer from serious conceptual as well as methodological defects (Lumsdaine, 1962e; Lumsdaine, 1963b), the fact is that they nonetheless do not provide convincing support for particular styles of programing in most instances (cf., Lumsdaine and May, 1965; Schramm, 1964b). Though they are not capable of logically showing that proposed principles are necessarily faulty, they do not offer sufficient evidence for putting forth such principles as bases for assessing programs in terms of their internal characteristics. A perhaps more direct form of evidence is supported in a study reported by Rothkopf (1963), in which individuals who had been trained in programing principles were asked to predict the relative effectiveness of seven different forms of a program, and their pooled and individual predictions were subsequently compared with the effects as actually determined by experimental measurement for these same program variants. The scope of this investigation was limited, and the programs studied were doubtless too short to exemplify the operation of all of the factors believed important in determining the effectiveness of programs. Nevertheless, the results of the comparison were far from reassuring. Not only were Rothkopf's "prophets" of effectiveness unable to predict correctly; their predictions showed a high *negative* correlation with measured effectiveness.

Such findings lend weight to the rational grounds suggested by Lumsdaine and Glaser for the importance of avoiding premature "freezing" of program styles (Lumsdaine and Glaser, 1960, p. 566). The evidence accumulated since that time has helped to illuminate some facets of the art of programing, but still falls far short of approaching a sufficient basis for any confident assessment of program effectiveness in terms of internal characteristics.

The need for test data to assess a program's effectiveness is summarized by the Joint Committee in the following excerpt from their 1962-63 report (AERA, 1963): "At the present time, the principal recommended use of internal data obtained from inspection of the programed materials is for determining whether program *content* is appropriate to the educator's objectives."

Of course one should not despair of the eventual possibility of accurately predicting program effects, and the firm validation of some predictive criteria by experimentation ultimately is to be expected. Even at present, it of course does not follow that all judgments would be as bad as those found by Rothkopf. In the long run, quite aside from the matter of efficiency in reducing the amount of trial and error needed to develop effective programs, validation of internal criteria even on a probabilistic basis is obviously desirable. This would permit making demonstrably valid estimates (even if only approximate ones) of possible effects prior to their being actually determined or verified by experimental measurement.

Tryout and Revision as a Basis for Gauging Effectiveness

The requirement of program tryout and revision is a central one in the programing rationale and has even been made a critical characteristic in Markle's definition (Ely, 1963). It seems obvious that one should be able to improve a program by testing its outcomes and progressively revising it until one has corrected the difficulties shown by the tryout test data. Is it possible, however, to use information about the developmental tryout and revision as a basis for assessing a program's effectiveness? Surely the mere fact that a program is reported to have been subjected to a tryout and revision procedure does not by itself assure that it has thereby become perfected. Lacking a validated, well-defined, and reproducible procedure that will demonstrably assure satisfactory results, validating data are still necessary for each program tryout and revision.

The Nature of Defensible Effectiveness-Assessment Standards

From the foregoing it is evident that the kind of "standards" advocated here are not standards for prescribing program content, construction, or style. Rather, it follows from the empirical orientation here adopted that the standards of concern are *standards of adequacy in the conduct (and reporting) of studies to determine program effects.* Standards for program "quality control" are thus standards for the quality control of program *data*. If one is to rely on empirical data to gauge program effectiveness, he needs to have assurance that the data afford a valid measure of the relevant effects actually produced by a program. The rest of this chapter is mainly concerned with the question of how such assurance can be provided.

Description of effects vs. "effectiveness standards": Program effects mean the changes in educational outcomes or attainments that can be shown to result from a program's use. The "effectiveness" of a program sometimes refers to the extent to which the program's effects are satisfactory in the light of the goals set for its use. Although one commonly speaks of assessing program effectiveness, experimental tests per se can only reveal a program's effects; whether these are satisfactory for a given purpose involves standards of judgment that cannot be dealt with more fully here.

Some agencies have thought to prescribe standards of minimum acceptable effectiveness in terms of test scores--such as the Air Training Command's "90-90" standard: "All programmed instruction packages (PIP's) will be designed to fulfill the terminal objectives to a 90% level for 90% of the students and therefore produce a mean test raw score of 90% minimum" (U. S. Air Force, 1962). In commenting on an earlier version of this paper, J. C. Flanagan has seriously questioned the wisdom of promulgating such "standards" at the present time; certainly they are indefensible, and even dangerous, without more-nearly absolute measures of attainment than the kinds of tests generally used to measure program effects.

The term "effectiveness" as used in this chapter implies only the question of determining what effects a program *is* capable of producing, rather than standards for deciding how effective it *ought* to be in order to be regarded as of acceptable effectiveness. A disposition to think of program-effectiveness data as descriptive rather than "evaluative" seems likely to avoid misunderstandings, especially if it is recognized that any description of program effects will inevitably be to some extent incomplete.

What Is a "Program"?

The placing of reliance for the determination of program effectiveness primarily on empirical evidence concerning what the program teaches carries with it logical implications for the definition of what should be called a "program." Often the use of the term *program* has been restricted by various writers to materials which have particular characteristics of format and sequencing that are believed to exemplify principles derived from behavioral science. This is coupled with an emphasis on criteria for assessment of program quality in terms of internal characteristics determinable on the basis of inspection. Among experienced programers (Lumsdaine and Glaser, 1960; Rothkopf, 1961; and the J.C.P.I. to the contrary notwithstanding), we hear such "in-group" characterizations as: "That's a program? Why, it's nothing but a series of copying frames!" Characterizations of this kind may well turn out to be cogent evaluations when eventually validated by appropriate data on the effects achieved by competing "real" programs. But at present, such statements seem to imply the existence of a basis for validation which does not as yet exist as much more than an article of faith. They have the status of hypothesis rather than of verified principle.

The attempt to restrict the use of the term "program" to materials exemplifying particular preconceptions about the value of alternative program forms or styles, of course, may be a useful heuristic in teaching students a particular technique of programing. However, no matter how well based in behavioral science such notions may seem to their proponents, they will not do at present, as has been emphasized, as a basis for demonstrating or establishing the actual merits or deficiencies of specific programs. Any proposed instructional vehicle ought to be allowed at least to enter a competition in which its merits can be demonstrated on the basis of impartial evidence of what it can do; it ought not to be barred from competition because "it's not really a program" in terms of failing to adhere to preconceived notions of what a "program" should look like. Restricting the field by definition can only be self-defeating, particularly at the present state of the art. Even if there is a good reason to believe that a "series of copying frames" is an inept style of programing, precluding such a sequence from the chance to demonstrate what it actually teaches can set up a restrictive situation which can inhibit creativity and lead to dogma rather than to either a science or technology of instruction.

Thus, rather than attempting to prejudge *what a program should be,* at the present time a very inclusive definition is necessary in terms of what a program can be. The following definition has been proposed elsewhere: "An instructional program is a vehicle which generates an essentially reproducible sequence of instructional events and accepts responsibility for efficiently accomplishing a specified change from a given range of initial competences or behavioral tendencies to a specified terminal range of competences or behavioral tendencies" (Lumsdaine, 1964, p. 385). This definition, with a minimum of restrictive connotations, can encompass most of the forms of programs that have been proposed under the "programed instruction" banner. The definition not only makes no particular theoretical presuppositions, but does not even require individually paced progress or overt responses by the learner as qualifications for inclusion as a program. The variety of program types and styles admitted includes individual learning programs differing in terms of such factors as use of larger or smaller steps and varying kinds or amounts of student response, as well as any combination of linear or alternative ("branching") pathways. It also includes within its compass "programs" designed for fixed-pace and group presentation, as well as individually paced programs. It thus admits to a competition for demonstrable effectiveness programs for group presentation by film, television, or other media, as long as the instructional sequence is substantially reproducible, and the program, of whatever nature, is assessible in terms of its demonstrable effects on students. A somewhat similar but slightly less inclusive definition, given by Susan Markle (Ely, 1963, p. 64), requires empirical development of material in order for it to qualify as a "program."

Application to individual and group-instruction programs: Focusing on external or validating criteria of program effects makes most of the following discussion equally applicable to all styles and forms of programs because it concentrates attention on the changes in behavior effected by the program, regardless of the nature of the program that effected the change. Some special problems of measurement, discussed in a later section, do arise for self-paced programs as a direct consequence of individual variations in instructional time which they permit. But the main aspects of assessment methodology apply just as much to assessment of instructional television, or filmed instructional programs for group presentation, as they do to assessing the effects of individually paced programed instruction (See also Lumsdaine, 1959). This generality, of course, does not apply to many of the proposed internal criteria, which relate to particular features of individually paced programed instruction following current patterns and which are dealt with in some of the proposed criteria, or checklists, for program assessment.

Program Assessment as a General Problem in Education

The broad definition of programs given above suggests also the realization that the basic problem in program assessment is not just that of assessing "programed-instructional" materials per se. This is only a facet of the total problem of being able to measure and predict the effects of all forms of instruction, whatever their nature, since it treats effectiveness in terms of "output" as related to "input" (initial competence), without regard for the processes or program characteristics whereby this gain from input to output is achieved. In terms of a hardware-system analogy, the concern is with how to assess the effectiveness of programs in producing a given output in relation to a given input, considering the program as a "black box." The internal workings need not be known for this purpose.

Major Aspects of Program-Effect Assessment

In this chapter "program-effectiveness assessment" is roughly synonymous with measuring the *effects* produced by a program under some observed procedure of use. By "effects" are meant changes which can be directly observed or inferred from recorded observations of changes in the behavior of students as a result of the use of the program. Such changes may include gains or changes in knowledge, skills (both verbal and psychomotor), attitudes, interests, or motivations as identified by specific kinds of behavior which such terms are intended to connote. Furthermore, "effects of a program" means changes which can be validly ascribed to the use of a program when other sources of influence have been ruled out by appropriate scientific procedures.

Purposes of Determining the Effects That a Given Program Can Produce

Two quite different purposes of testing the effects produced by a particular program need to be distinguished, because differences in procedures as well as in measurement techniques are sometimes appropriate to these different purposes. Perhaps the most important distinction to make here is between studies that assess a program's effects in order to provide (a) *a diagnostic basis for program revision* and improvement, i.e., a basis for empirical guidance of program development or (b) a *reportorial* basis for describing performance characteristics of a specified, completed program. Data for a completed product indicate to a teacher or other user the outcomes he can expect the product's use to achieve. (Data showing what effects were achieved by use of a program that is *not* available for general use seem to serve little current purpose, except to demonstrate that some difficult-to-achieve kinds of outcomes can in fact be produced by program use). The kinds of data appropriate to these two purposes have considerable overlap, however, though the uses may differ considerably, and the two purposes should not be confused.

For "diagnostic" purposes, the effects of a program on a *number of specific points* related to its objectives needs to be separately measured. These might include certain points of factual information and a variety of specific skills which it is desired to create. Here one is little interested in the total score: the relevant interest is in subscores for content units and even in what is learned on each specific test question or point. To achieve adequate stability of results, this of course requires a larger sample than to detect differences of the same magnitude in an over-all score.

Diagnostic subscores are not only of utility in the revision of a program; for a completed, published program they can indicate what specific additional instruction may be needed in order to achieve defined goals. Such data, for example, give the teacher guidance on what points of the subject matter need special attention in classroom instruction, as distinguished from those which can be achieved from the program. The usefulness of over-all total scores information, on the other hand, is mainly limited to indicating whether it is worthwhile to use the program at all.

Error Rate

"Error rate" is an *external* criterion in the sense defined above, but cannot, as such, be considered a *validating* criterion. Errors on unprompted criterion frames within a program come closer, and data for subtests on program units (which might include or consist of such frames) come closer still. An end-of-program test, preferably administered after an appreciable interval so as to measure retention freed from the cueing influence of immediate context (extended sequence-prompting effects), probably comes as close as is often likely to be practical for current program-construction practices, though still further delayed retention, transfer, and relearning (savings) tests deserve attention as further steps toward ultimate validation. A fairly low error rate may be and, at least for prevailing forms of nonbranching programs, probably often is a necessary condition for an effective program; but it is far from being a sufficient condition for effectiveness. Too low an error rate may, in fact, militate against optimum efficiency in many instances. It is easy to attain a low error rate by consistent overprompting or by a nonbranching or fixed sequence so slow as to produce very few errors by the least able of a highly heterogeneous population of learners that could hardly be of optimal efficiency for those at the able extreme of the distribution. In considering "error rate" as a datum, one should at least distinguish error rates of several types of items: prompted items in a linear sequence, items used to decide branching (whether prompted or not, review or otherwise), and terminal-behavior frames calling for unprompted performance of the to-be-learned behavior.

Whatever the usefulness of error rates (especially when classified by type of frame) may be for purposes of program revision, the uncritical use of over-all error rates, especially for entire programs or for heterogeneous sets of variously prompted and unprompted items, is by now largely discounted as a validating measure of program effectiveness. For other discussions or comments on the use of "error rates" or error counts as measures of program effectiveness, see the papers by Geis (1962) and Lumsdaine (1963d).

Response Data as a Basis for Revision

The point-by-point or step-by-step feedback to the programer (as well as to the student) has been widely hailed as a crucial and even defining feature of the programed learning approach. However, feedback to the programer does not necessarily require overt response in the program. Few would doubt that feedback from the constant responding as a student goes through a program is a valuable source of leads to the programer. But one of the things that the programed instruction field at first seemed slow to realize is that correct responses within the program do not necessarily mean that the terminal behavior (that is, posttest or retention performance) will be adequate. If one must depend on satisfactory posttest or retention performance as a necessary basis for assurance that the program is "working" properly, then it could be argued that there is no special virtue in having a record of overt responses within the program as a way of providing feedback to the programer. There are as yet few instances where the value of feedback to the programer as a basis for revision has been demonstrated in terms of improved test performance for programs thus revised as compared with concurrent parallel test scores for earlier versions of the same program. Examples are studies by Silberman and others (1964) in the field of individual programed instruction and by Gropper and Lumsdaine (1961) in the case of instructional television. The main value of frame-by-frame response data, obtained for a relatively small number of subjects, is to provide leads to the programer as to where (and perhaps how) to revise his program. Such data may be a helpful basis for suggesting certain revisions; they are not an adequate basis for validating the program's effectiveness.

Major Considerations Entering into Assessment of Program Effects

Some of the main considerations that apply in the conduct and reporting of studies of program effects can be considered under three main topics: (a) Consideration of *criterion measures* encompasses the characteristics of tests used as indices of what students can do or "are like" after the program, as compared with what they can do or "are like" before going through a program. This includes the definition of program objectives or potential outcomes and development of appropriate criterion tests reflecting attainment of these outcomes. (b) Under *utilization procedures and experimental design* must be considered procedures and arrangements for sampling and administering programs to defined samples of students under controlled and reproducible conditions, together with procedures using the above-identified criterion tests, particularly in terms of the control of extraprogram factors that may influence criterion attainment. The procedures used may be either those of a "laboratory" test or a "field" test; in either case controls must be introduced such that the data will reflect in a valid manner gains produced by the program, as distinct from other possible sources of responses on the test. (c) Although *reporting of program effects* cuts directly across the above two aspects, it is useful to consider reporting also as a separate category. Critical problems are how to obtain reporting in uniform terms, so that terminology has the same meaning to all users, how to insure soundness of reported data in terms of its reproducibility, and how to provide meaningful reporting intelligible to the prospective program user.

Criterion Measures

Behaviorally Stated Objectives

The problem of describing precisely what is to be taught and what it is, therefore, that is to be measured as an outcome of instruction clearly is niether new nor peculiar to programed instruction; it is a general problem of educational planning and evaluation. Contributions to the question of behavioral specification of instructional outcomes stem in considerable part from the work of Ralph Tyler (1950) at the University of Chicago, later reflected in Bloom's well-known Taxonomy (1956). Perhaps the most influential contribution growing out of the more recent concern with programed instruction has been Mager's book (1962) on the specification of instructional objectives.

Various aspects of the issues and problems involved in defining educational objectives are more fully treated in a series of papers by Lindvall, Krathwohl, Gagné, Glaser, and Reynolds, Tyler, and others in the recent volume edited by Lindvall (1964). The concepts of "task analysis" (cf., Gagné, 1964; Miller, 1962) have had an important influence in increasing the emphasis on need for precise description of the specific behaviors comprised by such objectives. (See also Deterline, 1964; Gotkin, 1963; Markle, 1964.)

A number of other writers (e.g., Gilpin, 1962; Lindvall, 1964; Lumsdaine, 1962b) have also stressed the conviction that relatively more attention needs to be given to defining objectives in relation to "what to teach" as contrasted with efforts to improve knowledge of "how to teach." The wiser definition of educational objectives need not remain solely a rational or judgmental matter, but may be aided by empirical data. In particular, it can be expected that improved knowledge of "what transfers to what" will give a better basis for identifying specific instructional objectives which demonstrably, rather than just as an article of faith, lead to the broader kinds of competencies and behaviors that can be agreed on in general terms as basic goals of education. The problems of such inquiry, basically a question of transfer of training, lie beyond the scope of the present paper.

Objectives and outcomes: The orientation of focusing on criteria of demonstrated effectiveness suggests the appropriateness of providing program assessment data in the form of stated "performance characteristics" which indicate what contribution the use of a particular program is actually capable of making toward the attainment of the specified instructional objectives. The specification of objectives may be done by the program producer and, separately, by the teacher, educational administrator, or other prospective program user. Some special questions arise from the fact that a given user's objectives may differ from those the program writer or publisher had in mind and also from the fact that in actual use a program may have effects that neither the producer nor the user necessarily had in mind in his original formulation of objectives (cf., Shettel, 1964). Part of the task for an ideally comprehensive program-assessment study is therefore to identify any likely relevant effects to which a program's use may lead--including but not necessarily limited to those proposed by the producer and user--and then to determine whether these possible effects are in fact produced by the use of the program.

Some Factors in the Design of Instruments for Measuring Instructional Outcomes

The following are illustrative of some major areas of concern that need to be considered in relation to criterion measures, or capability tests:

1. Definition of the universe of behaviors that constitutes the competence which the programer is trying to create or, in any case, with respect to which the evaluator is examining the accomplishments of the program. (The *"or"* implies that these could include definition of *any* behavioral outcomes which might ensue from the program, even though they were not all necessarily intentions of the programer.) These definitions should be accompanied by examples. However, such a definition is not a test itself; it is an analytic definition of the behavior that specifies what test items are relevant (cf., Flanagan, 1951; Glaser and Klaus, 1962; Lindvall, 1964).
2. The problem of item sampling and of how a specific test can be described, as completely as possible, as a sample of the universe of behaviors.
3. The formal or descriptive characteristics of the capability tests (e.g., the basic properties of reliability and validity, if this is applicable). Also, the test's origin: Is it an *ad hoc* test, a standardized test, or a mixture? May it be best sometimes to use standardized tests, but with certain items excluded? How should the level of performance be specified? As a related topic, can the outcome or objective of a program be defined without reference to the content of the program, especially when transfer is part of the educational objective?
4. The question of what *kind* of a criterion of program effectiveness to use: specifically, the question of time to attain some specified criterion level, as distinct from level achieved after going through the program, where performance time varies from subject to subject, so that time and achievement level are both dependent variables--as contrasted in the situation in which time is constant so that achievement level is the only dependent variable. This necessarily overlaps with the question of procedural design.

Sampling of test items: In assessing programs it is important to keep in mind the concept that an attainment test is, in general, only a sample of some total population of items that represent the criterion performance. But sometimes it is very difficult to define the relevant population of items in such a way that it does constitute the total teaching objective--and also so that, at least in principle, any number of independent samples can be drawn from this population so as to form equivalent tests. This is a problem which psychometricians seem not to have fully solved, perhaps because the kind of requirements they generally have do not require doing so.

What is needed is a definable population of potential items. It is not often feasible to enumerate all of them; however, two things might be used jointly as the bases for defining the population of items: first, one or more examples, and, second, some kind of "generating function." Given a generating rule that would define the population, and some illustrations that help show what it means, presumably a competent psychometrician can write any number of parallel test forms; one doesn't have to have all possible items constructed in advance and then sample from an already-written pool of test items. The importance of having a large pool of equivalent items from which successive samples can be drawn is made much more acute when tests are used repeatedly for the purpose of making successive determinations, in a branching program, as to whether a student is up to criterion or needs more instruction (see below).

Requirements for program-effectiveness tests vs. individual or student achievement tests: The problem of measurement usually faced by the psychometrician is primarily one of assessing individual differences. Hammock (1960) and Glaser (1963) have discussed some differences in the theoretical requirements for "norm-referenced" attainment tests for use in measuring individual differences vs. "criterion-referenced" tests designed to measure the effects of programs. Whereas for the former purpose economic constraints often require "objective" types of questions to permit electromechanical scoring for large numbers of papers that must be processed, this restriction need not be imposed for the relatively small number of subjects needed to assess a program. (See also Heath, 1962; Jacobs, 1962 concerning other aspects of the use of tests in relation to programed instruction).

Experimental Design

Utilization Conditions

In determining and reporting the effects produced by the use of a programed instruction package, conditions of utilization, in a classroom or in the use of a program for individual study, may obviously have considerable effect on its effectiveness. As emphasized in the J.C.P.I. 1962-63 report (AERA, 1963), it is recognized that programs will generally be used in conjunction with other instruction. However, unless the contribution of of a program's use to the student's knowledge or competence can be separated out from the contribution of the sources of instruction, there is no defensible way to tell what the program itself contributed. This involves the experimental control of extraneous sources of influence, so that gains in knowledge, skills, or behavioral tendencies associated with the use of a program can be validly defended as results of the program itself, rather than other concurrent or prior sources of influence.

The need to control for related causal factors will also depend on the question that is being asked--on what is to be assessed. As noted above, it is often necessary to know the effects of the program per se when used under conditions which, however specified, involve a minimum of other related concurrent instruction. Even though the contemplated later use of the program may actually be in conjunction with other instruction, it can be contended that relatively unambiguous information about what it can accomplish by itself, without supplementation, is more useful than uninterpretable information about gains produced by some unknown mixture of program effects and other unspecified influences. On the other hand, it may often be the case that the real question for assessment--that is, the question that reflects the decision to be made about a particular program's effectiveness--is what the program will contribute when used collaterally with other materials or procedure of instruction. This is not at all an impossible question to answer experimentally, though it does generally involve the need for specification as to what the "other" instruction is to be. The usefulness of such information about joint effects of a program and "other" instruction will obviously decrease as the "other" instruction departs from complete specifiability and reproducibility.

In this connection, the distinction between so-called "laboratory tests" versus "field tests" may need to be reexamined and sharpened. No doubt both of these kinds of program tests are needed. "Laboratory tests" of programs, that is program tests which are conducted under relatively describable, controllable, standardized conditions, are analogous to the "brake-horsepower" rating on an automobile, which is a useful statistic even though one also wants data on hill-climbing and other road tests. The latter may be thought of as analogous to the data obtained from field tests of the program--less exact, but taking into account factors not encountered in the laboratory (cf., Center for Programed Instruction, 1963; Glaser et al., 1963; Schramm, 1964a).

Utilization procedure vs. experimental design: The procedure for utilization of a program (i.e., the way it is to be used) sets limits on the experimental design that is appropriate for measuring the effects the program produces within that pattern of utilization. However, more than one experimental design may be employed to determine the effects that are produced by a program when used in a particular pattern of utilization. Conversely, the same general design may be used with appropriate modifications for determining a program's use under more than one condition of utilization.. For example, either a "before-after" or "after-only" design can be used to determine a program's effects under a given condition of use, and either basic design can be used to compare a program's effects under two conditions of use. (After-only is used here in the special sense employed by Hovland, Lumsdaine, and Sheffield (1949), where a nonexposed control group's score, rather than preexposure scores for an experimental group, furnish the "before" measure.) But some of the important aspects of experimental design can be considered only in relation to the features of certain patterns or procedures of program use such as those distinguished below.

Effects as gains due to programs: Essentially, the question of measurement of program effects follows a "before-to-after" change paradigm. That is, the determination of the effects of any given program reduces to the problems of determining what the learners are like in terms of behavioral competencies and tendencies *following* the use of a program (or some segment thereof) as compared to what they would have been like had they not used the program, or what they would have been like had they used some specified alternative program. The same argument applies whether the program in question is a *total* program, as packaged for sale or distribution, or a program segment. The principal departure from this kind of model will be found in the case where program effectiveness or efficiency is measured in terms of time required by a person to reach a given criterion of accomplishment, rather than measured by achievement at the end of a fixed sequence of program materials.

Use of statistical significance tests: A weakness of the statistical habits associated with before-after and gain experiments is that the statistical tests employed are addressed to hypothesis testing rather than to estimation (cf., Schutz, et al., 1962). It is true that in determining the effects of a program, one wants to rule out the null hypothesis that observed gains can be dismissed as chance differences; i.e., one wants to show that effects produced were statistically reliable. However, what is obviously of more interest is a good estimate of the *size* of the gain; merely showing reliable evidence for *some* gain can be trivial. Unfortunately, as Lumsdaine and May have noted (1965, p. 490), the practice of reporting the size of effects on an interpretable scale and with accompanying confidence intervals is as yet more the exception than the rule in experiments of the effects of programs.

Some Factors in Experimental Design as Related to Program Assessment

Some of the problems of experimental design are independent of measurement problems per se; others interact with the kind of criteria that are employed. Several recurrent questions are: (a) The experimental procedure--that is, how to specify what was done in using the program, the instructions under which it was used (including the test instructions), how long people worked at it, under what kind of supervision, with what kind of incentive factors, etc. (b) Questions of control (not necessarily implying a control group) for extraneous sources of influence. (c) The need for alternative or complementary evaluation procedures to answer two different questions, previously identified: (i) What do subjects learn when some form of control insures that they do, in some sense, go through the program? (ii) And if a program is merely made available to people, to what extent *will* they go through it (and learn from it)? In other words, to what extent is the program self-motivating? A related design problem is that of how to take account, in analysis and reporting, of dropouts which occur even under a relatively controlled situation. (d) Differences in the kinds of requirements that would apply for experimental design in the case of relatively *brief* programs as compared with quite lengthy programs.

In any study on assessing "effects" of a program as gain from "before" to "after," several methodological decisions must be made. If there is good reason not to be worried about the effect of concurrent extraneous events, it is possible to simply measure the same group before and after using the program. Thus less test data would probably be needed in the case of very short programs. But generally, with any lengthy program it is necessary to introduce a suitable control for what is happening in the meantime: what the students are eading on the outside, what help their parents ar giving them, and so forth. This in turn generally means that to assess the effects of a long program, some form of control group is necessary; simply a before-and-after measure for one group will not suffice. Note that such a control group does not mean a group that was given "conventional instruction" as a base of comparison; rather it means a "nonexposure" control for extraneous influences so as to afford a measure of what the program group would have been like had they not received the program. But such a control suffices only for extraneous influences that act independently of the program's use; for other influences (e.g., outside help given on the program itself), other control strategies are needed.

The extent to which formal requirements of control actually are of practical importance in a particular assessment situation often can be tempered by the judgment of the experienced experimentalist. For example, the degree to which pre-to-post changes for a nonexposure control group must be subtracted from those found for the experimental group will depend on the likelihood that substantial changes will occur as a result of extraneous (nonprogram) influences. This likelihood will be much greater in some cases than in others.

One extreme might be the case of a semester-long program on a topic which figures largely in current discussion and public news. It is conceivable in such cases that *all* the gain observed from before-to-after a program might derive from outside sources and the program itself be wholly ineffective. Here the need for a control group not exposed to the program is more obvious than, to take an opposite extreme, in the case of a one-half hour program which can be given and tested without students leaving the experimental classroom, thus affording nearly complete control of extraneous informational sources. (For further discussion of alternative experimental designs for determining instructional effects, see Campbell and Stanley, 1963; Lumsdaine, 1963b).

Reporting of Program Effects

Need for Technical Standards in Measuring Program Effects

A need to provide guidelines for the consistent and interpretable reporting of tests to assess the achievement of programs conmplements the need for standards to guide the conduct of such tests. The idea of technical recommendations for the assessment of programs in terms of what they demonstrably teach essentially implies some form of controlled experimental measurement which, as indicated by the 1962-63 Joint Committee report (AERA, 1963), can yield evidence to "document for the technical reader that the gains in achievement reported can rightly be attributed to the effects of the program's use rather than to extraneous causes." It is assumed that data on the effectiveness of programs will be obtained and reported by various agencies, including program producers, using agencies (including school systems), and projects conducted by universities and other research agencies. Three levels of reporting may be distinguished: (a) Summary reports, advertisements, general characterizations of program effects and uses; (b) teachers manuals, giving details of program effects demonstrated under described conditions of use, in sufficient detai. that valid results may furnish a usable guide to selection and use of programs; (c) technical reports amplifying the teachers-manual information in the kind of detail needed for a technical expert or consultant to assess the validity and applicability of the data summarized in the teachers manual. (This could be a technical supplement to the teachers manual, or, since fewer copies will be needed, a separate technical report.)

Reporting Descriptive Information for Program Users

It seems very desirable to provide user's manuals (or teachers manuals) for programs, as a vehicle for presenting relevant external information about properties which are not apparent on inspection. Information presented in a manual might deal with (a) the program's purposes and intended use, (b) the source of program content, (c) the way the program was developed, including tryout and revision, and (d) the conduct and results of testing to determine empirically the effectiveness, or "performance characteristics," of the published program. The last of these kinds of information is, of course, the most relevant to the present discussion.

Reporting of Information About the Demonstrated Effectiveness of a Program

It is to be hoped that manuals for programs, at least for programs of considerable scope, will furnish evidence on the program's effectiveness based on measurement of student performance on pre- and postprogram criterion tests. The J.C.P.I. has recommended that these tests be exhibited either in the manual or in a supporting technical report, so as to exemplify what the producer expects students to learn as a result of program use. Suggested content of information for teachers manuals concerning reported program effects is presented in a supplement to the 1964 J.C.P.I. report (AERA, 1964; National Education Association, 1964). Information on effects, presented in nontechnical terms heopful to the teacher or program purchaser, will generally need to be backed up by a more detailed technical supplement to permit technical assessment of the adequacy of the data presented. Outcomes resulting from the use of a program need to be described as concisely, objectively, and simply as possible, with the aim of communicating to teachers and supervisors how the program was used and what results this use produced.

The writer believes that the emphasis in the reporting of assessment data should be primarily descriptive rather than "evaluative" in the sense of passing judgment on the desirability of different kinds of objectives and outcomes. The aim of such data is to provide a clear picture of what each program *will do* under two or more conditions of use, rather than to pass judgment on what the program *should do*. This leaves it to schools to decise whether the kinds of outcomes that can be realized are the ones which they wish to attain. However, relevant normative data are obviously useful for comparison where available and appropriate.

Description of Effects for Published Programs

The writer believes that claims for the effectiveness of a published program should be supported by data from evidence of gains in student attainment produced by the *final, published version* of the program. A clear distinction should be made between this "effectiveness-test" data, for the final program, and any test data obtained in earlier tryouts of preliminary versions, used as a basis for revision of the program. The sole purpose of these earlier data is to point the way to program revision. By contrast, any changes made in the program after "effectiveness" data are obtained could throw doubt on the validity of these latter data for furnishing a demonstration of the program's effectiveness, since changing the program could lessen as well as increase its effectiveness. The prospective user should have data, if possible, based on the edition he is trying. If data based on an earlier edition are offered in support of a program, this fact should at least be clearly stated, since at the present state of the art there is no real guarantee that revision, particularly if based on editorial judgment, will necessarily have improved a program. In at least one case known to the writer, editorial revisions almost certainly reduced the quality of the program and thus impugned the validity of test data offered in support of it.

Standards in Reporting Findings of Educational Research

Assessment of purportedly factual statements in the literature about education is very difficult. For example, there is no way to assess the significance of statements such as "average students chose televised instruction 6 to 4 over face-to-face instruction" or "more questions are being asked than in classes taught by traditional methods." There are no standards of consistent reporting, no assurance of dependability, such as one would have in a summary abstract or a technical journal in physics, or even, to some extent, in Psychological Abstracts. This is not entirely just a difference in the rigor and precision obtainable by science, but is partly a question of the *recognition* of the existence of a technical literature, of technical bases for making statements, and of conventions about what is said on the basis of what kind of evidence--or rather, what kind of evidence can reasonably be presumed to lie behind what kind of statement. One can expect loose statements in the press, but is often not so wary in reading a professional journal or book. The problem is that the line between journalism and technical reporting is much hazier in the field of education. Perhaps it may be possible to establish standards of consistent reporting or of evaluation of published statements which would reflect the kind of evidence which actually lies behind a given statement. A real inquiry into the possibilities would be a worthwhile effort. The importance of this problem will increase as more and more educational literature accumulates which tries to report scientific data about the effects of educational programs.

In relation to the problems of reporting program effectiveness, it should be noted that impressions of program effects often are not clearly differentiated from actual findings documented directly by quantitative evidence. Yet, the former, by virtue of the way they are stated, are likely to be given as much credence as documented findings. This seems particularly likely when conclusions based on mere impression are presented in the context of an actual research study in which certain data were found which do justify certain conclusions. The former thereby acquire an aura of validity ("gilt" by association?) which they do not, in fact, possess.

The J.C.P.I. (AERA, 1964) strongly recommends that summary reports describing a program's effects (in the form of press releases, advertising, or teachers manuals) be withheld until there is also available a technical report setting forth the procedures by which the data were obtained. Such technical information, in sufficient detail that the technically qualified adviser may be able to assess its soundness, should be available to back up claims made in summaries prepared for administrators or teachers. It is important that this basis be provided in order that data which are actually sound and valid indicators of the effects of a program will not be confused with those popularly presented data which, in fact, do not validly reflect the effects of a program.

Reproducibility

A general criterion of the value of criteria for reporting as well as conducting any assessment study is that the evaluation procedure and its results should be reproducible. This applies to the derivation, administration, and description of criterion measures as well as to the selection of the experimental population and all aspects of the design and procedure. The object of the reporting is to describe what the program accomplishes, in such a way that the process can be duplicated with substantially similar results. The reporting of evaluation tests should therefore describe the physical and social conditions of the program's use and the effectiveness-testing procedures in sufficient detail so that their essential features can be reproduced by another investigator if desired. Any discrepancies between recommended conditions of use and those that were employed in obtaining the effectiveness-test data should be noted.

Technical reports should indicate how many students started and completed the program, average completion time, the average level of performance on the specified pre- and postprogram tests of achievement, and the range or variability of these measures. Relevant further temporal data would include the amount of time learners of different ability spent on various portions of the program, how this time was distributed (especially for long programs), and how long after completion of the program the learners were given the criterion test (See also AERA, 1964; Rothkopf, 1961).

The AERA Joint Committee General Recommendations on Reporting

The J.C.P.I. (1964) makes the following three general recommendations about the reporting of evidence on a program's effectiveness:

First: *Evidence for the effectiveness of a program should be based on a carefully conducted study which shows what the program's use accomplished under specified conditions.*

Second: *The results of the evaluation study should be carefully documented in a technical report prepared in keeping with accepted standards of scientific report.*

Third: *All claims or statements about the effectiveness of a program should be supported by specific reference to the evidence contained in the technical report.*

The substance of the Committee's further recommendations for prospective purchasers or users of programs includes, in addition to the general caution of evaluating each program on its own merits, the following further recommendations: (a) That in determining the usefulness of any particular program, a prospective user should first try to formulate his own objectives in as much detail as possible and then evaluate the program, in relation to these objectives, in the light of the three criteria of suitability, practicality, and effectiveness in attaining outcomes relevant to his own objectives. (b) That the prospective user should ignore all claims for the effectiveness of a program which are not backed up by appropriate data that have been subjected to competent evaluation. It is further recommended that advice on the soundness of claims for program effectiveness preferably be obtained from a technical adviser who has competence in the fields of educational psychology, measurement, and experimental design. And (c) that users should seek all available data on the demonstrated characteristics of the program, both from information supplied by the producer and also from reports prepared by school systems, research projects, or other agencies that have conducted program-assessment studies of the particular program.

Additional recommendations for program publishers are proposed by the Committee to assist them in providing necessary information which will help the user make an intelligent choice of a program. These include the recommendations that (a) publishers state in detail the objectives of each program, preferably in terms of specific behavior or competence which its use is intended to achieve; (b) publishers cite the available evidence to document the statements they make about the effectiveness of the program, citing any pertinent evidence available both from their own studies and from other appropriate studies of the program; (c) publishers refrain from promoting a program on the basis of unsupported statements about its effectiveness or in terms of general statements about the value of the programed instruction "method"; (d) publishers provide a program manual, preferably one that can be updated or supplemented as new data on the program become available, and one that substantiates all claims made in this manual or elsewhere by citing documentary evidence from a technical report of a carefully conducted evaluation study. Publishers are also advised to differentiate explictly between opinions about the effectiveness of the program and documented evidence on the outcomes it can be shown to produce.

Technical consultation: It is recognized that the teacher or administrator will generally not wish to be burdened with the detailed technical information that is necessary to provide an adequate basis for determining the soundness of reported experimental findings. Consequently, data put forth purporting to indicate a program's accomplishments generally cannot be validly interpreted except with the consultation of a technical specialist who has examined the procedures and instruments whereby the data were obtained. The requirement for having such technical consultation available to appraise data on the effectiveness of instructional programs is a relatively new one in educational institutions. However, the utilization of other kinds of technical experts as consultants in educational decisions, for example, engineering and architectural consultants, is, of course, commonplace. As educational technology advances, it may be expected that a similar need may be more widely recognized for technical specialists to advise the educational administrator and curriculum supervisor on the validity of data about the effects of educational programs, just as he now calls on experts to advise him with respect to the characteristics of audiovisual equipment or the construction of instructional facilities.

Techniques for describing program effects: Describing program effects in a useful way for teachers and administrators cannot be done by requiring them to pore over detailed tabulations of data. Simplified techniques of presentation are needed, yet must be reconciled with the fact that the data are basically rather complex. For example, in describing the performance of students on a program, a "three-dimensional" distribution is needed. Each person started at some level and got to a different level in a certain amount of time. There are, therefore, three descriptive dimensions: a starting point, a terminal point, and the time required to get there. There is also a need to present subscores as well as total scores, and for both there will often be considerable variation among learners in their starting points as well as variations in the terminal points or levels reached. In addition, individual differences in learning ability also lead to variations in how long it takes students to get to any specified level in using the program. Since each of these variables involves a frequency distribution, a quite complicated descriptive problem exists.

Various ways of attempting to cope with this problem can be seen. In reporting gains in total score from pre- to posttests, evaluation studies by publishers (Basic Systems Program, 1962a, 1962b; Drooyan and Wooton, 1964) have simply tabulated the pre- and posttest scores individually for each subject for whom data are reported, with N's ranging from 3 up to 30 or more. Time spent on the program by each student is similarly reported. Glaser, Reynolds, and Fullick (1963) present graphical distirubtions of scores on the pre- and posttests rather than tabulating individuals' scores. A recent paper by Hively (1964) and graphic presentation techniques used by Hovland, Lumsdaine, and Sheffield (1949) and Zuckerman and Jacobs (1951) also suggest several graphical methods of presentation. None of these, however, has attempted to integrate the time and gain variables.

Some Methodological Problems in Improving Program-Assessment Practices

Various difficulties of a methodological sort are encountered in current program-assessment studies. Several key problems are identified below. (See additional discussion of some of these problems in an earlier paper [Lumsdaine, 1963d].)

Learner characteristics: Specification of prior knowledge and ability of learners can serve both to identify the preprogram baseline from which gains may be measured and also to indicate what prerequisites are needed in order to learn effectively from the program. The corresponding characteristics for the samples of students used in preliminary tryouts or, particularly, in the effectiveness testing of the program, need to be separately specified so as to indicate the degree to which these learners were typical or atypical of the learners for whom the program is intended. A special problem here that calls for some methodological research is that of determining "latent" initial knowledge, based on prior knowledge not recalled, but readily relearned, which is thus not revealed by a preinstructional achievement test.

Problems in determining initial level or "entering behavior": There are two kinds of initial capabilities from which gain to terminal behavior capability takes place. The first is the degree of initial competence on the specific set of behaviors which are to be modified by instruction. The second comprises other, related initial capabilities which are not to be modified, but which are assumed to be prerequisites for the desired modification which the program is to accomplish. For example, in addition to "an initial and a terminal capability" in spelling or in ability to perform algebraic manipulations, an initial capability (not to be modified by the program as such) in reading ability and perhaps in competence in writing (or button pushing?), if the program calls for these kinds of responses, is also assumed. Some general intellectual competence, summarized under "IQ" or described by age level, school placement, etc., often without further analysis, is also assumed. The attempt to analyze fully and completely all the prerequisite capabilities would be a very demanding task. On the other hand, being as complete as is feasible is likely to pay off, since it may point to supplementary program requirements to augment additional prerequisite capabilities not initially thought of.

Adequate identification of initial capability is one of several problems encountered in assessment of program effects which suggest inherent deficiencies in programing rationale. The straight-through linear program not only does not lend itself well to measuring time to criterion (see below), it also invites a spuriously inflated gain if initial knowledge is present that is not revealed by a pretest. These two difficulties are interrelated; the latter one can be minimized if a brief review is given between two pretests before starting the main program. But a more fundamental solution is to change the nature of the program itself by branching sequences that skip over detailed coverage of the material on which an initial overview or brush-up has already shown satisfactory achievement.

<u>Time vs. criterion achievement</u>: A special problem is brought into focus by the assessment of self-instructional programs, where time spent in instruction is a dependent variable as well as gain in achievement level. In dealing only with fixed-pace programs such as films and television, one could readily compare, say, two versions of a film, each of which took one hour to show. In this case, the problem of evaluating time versus achievement was not as apparent. If one film produced higher achievement scores, and it could be demonstrated that this superiority was reliable,

it could be said that it was the better film. The comparison of two self-instructional programs, one of which produces higher achievement scores than the other but also requires more time for students to finish is much more difficult. Of course, it is possible simply to report two separate sets of facts. This is perhaps about the best solution at present; it is not a very good long-range solution because it leads to no decisive basis for perferring one program over another when one program scores better in terms of achievement, but an alternate program scores better in terms of time. Gain in achievement level has sometimes been expressed as an "efficiency" ratio of gain divided by time (Goldbeck and Campbell, 1962). Objections to this procedure stem from such considerations as nonlinearity of the achievement-gain scales. These considerations, however, do not controvert the need to take time into account in some such fashion.

At present, a considerable advance in program description will be made if *any* objective records are presented that clearly identify the intial and terminal points and the gains that can be unambiguously attributed to the effect of using the program. As a complement to these "achievement-gain" data, it would be desirable to see the time-in-study data even if, for the present, no single achievement-time index seems defensible as a single figure of merit for a program's instructional efficiency. In describing the effects produced by the use of a program, any of the following may be useful (see Lumsdaine, 1963d):

1. Report gains in attainment of outcomes achieved by going through the program from beginning to end and separately report time spent on the program as a second, separate dependent variable.
2. Determine and report as the main dependent variable time required to achieve specified levels of attainment.
3. Hold time constant, reporting attainment achieved in some arbitrarily fixed period of time.
4. Let both time and attainment vary, using some devised single measure such as amount of attainment per unit time.

Time to reach a criterion: When agreed-upon minimum levels of proficiency to be attained can be set, the time that is required to get to these levels can be used as a criterion instead of comparisons between a gain from a given before-level to a given after-level (cf., Gilpin, 1962; Skinner, 1958). In order to know more or less continuously whether a learner has reached the criterion or how closely he has approached it, one has to use frequent interspersed tests; certainly one cannot give the whole program and then measure at the end of it. But, in order not to spend all of the student's time in being tested, this requires some form of sequential-sampling test, in which one or two responses are used as the basis for a branching decision--that is, the decision of whether to stop testing and to proceed with further instruction, or ask more questions, as determined on the basis of the first question or two.

"Savings" measures: There is a potential alternative to measuring time to criterion. This alternative is to use as a measure the time required to get the student up to (or back up to) some criterion by further instruction of a specifiable sort. If the student had initially achieved this criterion, such a "savings" measure is the time spent in *re*learning, as in the traditional use of "savings" measures of retention in the psychological laboratory.

Savings measures have the advantage of permitting translation of program effectiveness into time values that can be evaluated in terms of cost and yet which can be applied as an indication of retention at any time after a student has finished a program. Moreover, even if the student never attained criterion performance by the time he initially completed a program, a "savings" measure can still be used, indicating how much more instruction was required to get him initially up to a criterion level of performance. (Variants of the basic idea of using "savings" measures for assessing program effects seem to have been independently arrived at by R. Glaser and E. Z. Rothkopf, in discussion with the present writer.

Retention: In using retention as a criterion, the interest is not primarily, of course, in just how well a student does at the end of any given lesson or even in how well he does at the end of the semester. Furthermore, it may be of less interest to know whether he can get a perfect score on a recall test six months (or two years) after he has finished algebra than to know how fast he can relearn algebra when he needs it (and how fast he can relearn it with what kinds of refresher instruction). For many kinds of learning this is the criterion on which more and more premium will have to be placed as knowledge continues to expand. Thus, to assess the effectiveness of a program properly, the concern sooner or later must be not only with whether it gets a person up to some immediate criterion, but with how well it compares with some other program (that may have gone up faster or slower to perfect immediate performance) in terms of a "savings" score. The principal problem in the use of a "savings" measure, whether used to measure immediate or delayed retention, is that of specifying a standard or reproducible vehicle for providing the "refresher" or "finishing" instruction. Nevertheless, the "savings" approach is so attractive that considerable effort appears warranted to solve the methodological obstacles it presents.

Empiricism and Theory: Today and Tomorrow

Many challenging problems are posed for research needed to improve the methods of instructional program assessment. One of these, noted above, is the problem of how to predict, from intraprogram response data, what the long-term recall or relearning performance will be at a later date. Better solutions to this and other methodological problems need to be found, though optimum solutions are not likely to be obtained in the immediate future. For the present, however, useful work in assessing program effects can still be done with methods which fall far short of those that hopefully will be available five or ten years hence.

Many statements have been made about properties of programs regarded as related to their effectiveness, and a number of experiments have been conducted to test the predictive validity of such propositions. The fact that some reasonable propositions about the advantages of theoretically sound prompting techniques and sequencing of content are far from fully supported by available evidence does not necessarily mean that such propositions are without validity nor that they should necessarily be ignored as guides to programing. While this may be obvious, it seems well to make it explicit in view of the stress placed here on the need for reliance on empirical evidence to assess each program's effectiveness. This empirical emphasis does not deny the usefulness of theory as a basis of prediction or a guide to program construction. It merely takes the position that for the present, theoretical propositions and rules of programing derived from them are insufficient and therefore hazardous bases on which to rely for the assessment of a program's effectiveness. Obviously it would be more efficient as well as more elegant if results of empirical tests of effectiveness could be reliably predicted in advance by analysis of program construction features, i.e., on the basis of "internal criteria." It is certainly to be hoped that this will be the case in the future. At present, it is to be recognized that a mature science of instruction does not exist, and wishing won't make it so. How soon research will make it so remains to be seen.

In the meantime, as Candide had it, "We must cultivate our gardens." The weeds to be gotten out of the programed-instruction garden are generalizations about "the method" based on faith rather than evidence, unsupported claims about program effectiveness, and data advanced for such claims which, however extensive or neatly presented, actually fail to support the claims in terms of accepted standards of scientific evidence. We need to rid the programing field as rapidly as possible of these impediments to its orderly growth, and thus put on a sound basis of practice the philosophy of providing tested performance specifications for current programs. We may then profitably shift more of our research emphasis to the longer-range goals of cultivating the knowledges and understanding that will be needed to comprise a science of curriculum, which can afford a rational basis for deciding on what programs need to teach, and a science of instruction, which can more dependably guide our efforts to teach it efficiently.

Footnotes

1. The preparation of this paper was materially aided by research projects supported by the Ford Foundation and by the Educational Media program of the U. S. Office of Education, including support by the latter for the work of the AERA-APA-DAVI Joint Committee on Programed Instruction (see also footnote 2). Reproduction in whole or in part is authorized for any purpose of the U. S. Government. The writer, who serves as chairman of the J.C.P.I., would like to acknowledge the contributions of his colleagues on the Committee (H. F. Silberman, E. R. Keislar, and Robert Glaser [AERA]; R. S. Crutchfield, J. G. Holland, and L. M. Stolurow [APA]; and J. V. Edling, E. B. Fry, W. C. Meierhenry, and P. R. Wendt [DAVI]). He also wishes to thank E. Z. Rothkopf, L. C. Silvern, B. B. Hamilton, and other consultants and staff assistants for their assistance in the Committee's work, to which several portions of the present paper are closely related. The author is indebted to J. C. Flanagan for helpful comments on an earlier version of the paper and to Harriet Foster, Susan Markle, M. J. Rosen, and other colleagues and students at UCLA for their contributions to this notions on program assessment. It is not, of course, implied that any of these individuals shares all of the opinions stated in the present paper or bears responsibility for any deficiencies of content or exposition which it contains.

2. Most of the work of the Committee has been supported, under the provisions of Title VII of the National Defense Education Act, through a contract from the Educational Media Branch of the U. S. Office of Education to the American Educational Research Association. For brevity this committee is referred to herein simply as the "AERA Joint Committee," as well as by the abbreviation "J.C.P.I." The Committee's activities through 1962 were summarized in an article by the chairman (Lumsdaine, 1962a), and its published reports are cited in a number of places in this chapter.

3. For definitions of technical terms used in describing program features, see Ely, 1963; Markle, 1964.

4. Among these guidance statements and checklists, in addition to the previously mentioned statements by Rothkopf (1961) and by the AERA Joint Committee (1961, 1963), have been offered by Belton (1962), Jacobs and others (1964) Tracey (1963), the USAF's Air Training Command (1962), the New York Board of Education (1962), the Rocky Mountain School Study Committee (1962), the Center for Programed Instruction (1962), the University of Michigan Center for Programed Learning for Business (1963), and the National Society for Programmed Instruction (1962, 1963). Guidance statements or checklists have also been provided by several of the commercial program producers, including TMI-Grolier (Teaching Materials Corporation, 1962), Coronet (Coronet Learning Programs, n.d.), and the General Programmed Teaching Corporation (1963). Internal criteria are also implied in various guides to programing procedure--e.g., Klaus (1961) and Wiley (1961)--as well as in several textbooks (cf., Lumsdaine and May, 1965, p. 480).

References

American Educational Research Association, American Psychological Association, Department of Audiovisual Instruction, National Education Association: Joint Committee on Programed Instruction and Teaching Machines. Joint committee prepares interim guidelines. AV Communication Review, July-August, 1961, 9, 206-208.

American Educational Research Association, American Psychological Association, Department of Audiovisual Instruction, National Education Association, Joint Committee on Programed Instruction and Teaching Machines. Criteria for assessing programed instructional material. Audiovisual Instruction, February, 1963, 8, 84-89.

American Educational Research Association, American Psychological Association, Department of Audiovisual Instruction, National Education Association: Joint Committee on Programed Instruction and Teaching Machines. Recommendations for reporting of information on the performance characteristics of programmed learning materials. Third interim report. Los Angeles: University of California, 1964. (Preliminary edition)

Basic Systems Program. Chemistry I, Atomic structure and bonding: Teacher's manual. New York: Appleton-Century-Crofts--Lyons & Carnahan, 1962a.

Basic Systems Program. Vectors, A programmed text for introductory physics: Teacher's manual. New York: Appleton-Century-Crofts--Lyons & Carnahan, 1962b.

Belton, J. R. Programmed learning--20 questions. Bulletin of the National Association of Secondary School Principals, December, 1962, 46, 77-78.

Bloom, B. S. (ed.) Taxonomy of educational objectives. Handbook I: Cognitive domain. New York: Longmans, 1956.

Bolt, Beranek, and Newman, Inc. Test results of honor roll programs. Brighton, Massachusetts: the Author, 1963.

Buros, O. K. (ed.) The fourth mental measurements yearbook. Highland Park, New Jersey: Gryphon Press, 1959a.

Buros, O. K. (ed.) The fifth mental measurements yearbook. Highland Park, New Jersey: Gryphon Press, 1959b.

Campbell, D. T., and Stanley, J. G. Experimental and quasi-experimental designs for research on teaching. Handbook of research on teaching (edited by N. L. Gage). Chicago: Rand McNally, 1963, 171-246.

Carroll, J. B. Sorcerer's apprentice at large? Contemporary Psychology, April, 1964, 9, 188-90.

Caulfield, R. L. Evaluation of program texts. School and Society, March 7, 1963, 91, 116-117.

Center for Programed Instruction, Information Division. A parents' guide to teaching machines and programed instruction. New York: the Center, 1962.

Center for Programed Instruction. The use of programed instruction in U.S. schools. New York: the Center, 1963.

Coronet Learning Programs. How coronet learning programs are prepared and tested. Chicago: Coronet Instructional Films, n.d.

Denova, C. C. Byways through a FORTRAN maze. Contemporary Psychology, November, 1963, 8, 456-58.

Deterline, W. A. The carpentry revolution--a fable. Trends in programmed instruction. (Edited by G. D. Ofiesh and W. C. Meierhenry.) Washington, D.C.: National Education Association and National Society for Programmed Instruction, 1964, p. 11.

Drooyan, I., and Wooton, W. Teacher's manual to accompany programmed beginning algebra. New York: John Wiley & Sons, 1964.

Edgerton, H. A. The acceptability and effectiveness of the casual use of auditory training aids. ONR Technical Report No. NAVTRADEVCEN 373-1. Port Washington, New York: U. S. Naval Training Device Center, March, 1960.

Eigen, L. D. Scrambled bits. Contemporary Psychology, January, 1961, 6, 25-27.

Eigen, L. D. Some problems in field testing programs for teaching machines. Journal of Educational Sociology, April, 1961, 34, 372-76.

Eigen, L. D. The implications for research methodology of some behavioral studies in programed instruction. Psychology in the Schools, April, 1964, 1, 140-47.

Ely, D. P. (ed.) The changing role of the audiovisual process in education: A definition and a glossary of related terms. AV Communication Review, January-February, 1963, 11, 1-148.

Esbensen, T. The programmer as artist. Auto-Instructional Devices, November, 1962, 2, 186-87.

Flanagan, J. C. The use of comprehensive rationales in test development. Educational and Psychological Measurement, Spring, 1951, 11, 151-55.

Fletcher, L. The non-programmed program. NSPI Journal, May, 1964, 3, 6-7.

Gagné, R. M. The implications of instructional objectives for learning. Defining educational objectives (edited by C. M. Lindvall). Pittsburgh: University of Pittsburgh Press, 1964, pp. 37-46.

Galanter, E. Common statistics uncommonly taught. Contemporary Psychology, March, 1961, 6, 104-106.

Geis, G. L. Some considerations in the evaluation of programs. AV Communication Review, January-February, 1962, 10, 64-69.

General Programmed Teaching Corporation. Judging program quality. Guide to programmed teaching. Albuquerque, New Mexico: the Corporation, 1963, pp. 15-18.

Gilpin, J. Design and evaluation of instructional systems. AV Communication Review, January-February, 1962, 10, 75-84.

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

Glaser, R., and Klaus, D. J. Proficiency measurement: Assessing human performance. Psychological principles in system development. (Edited by R. M. Gagné). New York: Holt, 1962, pp. 419-74.

Glaser, R.; Reynolds, J. H.; and Fullick, M. Programmed instruction in the intact classroom. Pittsburgh: Learning Research and Development Center, University of Pittsburgh, December, 1963.

Goldbeck, R. A., and Campbell, V. N. The effects of response mode and response difficulty on programmed learning. Journal of Educational Psychology, June, 1962, 53, 110-18.

Gotkin, L. G. Some guidelines for the location, selection, and individual testing of programed instruction for school use. Paper read at the National Conference on Selection and Use of Programed Instruction, Detroit, June, 1963.

Gropper, G. L., and Lumsdaine, A. A. The use of student response to improve televised instruction: An overview. Studies in Televised Instruction, Report No. 7. Pittsburgh: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX and the American Institute for Research, June, 1961.

Hammock, J. Criterion measures: Instruction vs. selection research. Paper presented at the Annual Convention of the American Psychological Association, Chicago, September, 1960.

Hanson, L. F. (ed.) Programs, '63: A guide to programed instructional materials available to educators by September 1963. Washington, D.C.: Government Printing Office, 1963.

Heath, R. W. Some uses of tests in programmed instruction. Nineteenth yearbook of the national council on measurement in education (edited by M. Katz). Ames, Iowa: National Council on Measurement in Education, 1962, pp. 46-48.

Hendershot, C. H. Programmed learning: A bibliography of programs and presentation devices. Second edition. University Center, Michigan: Delta College, 1963.

Hively, W. Defining criterion behavior for programmed instruction in elementary mathematics. Cambridge, Massachusetts: Committee on Programmed Instruction, Harvard University, 1963.

Hively, W. Some guidelines for evaluating programs of instruction in mathematics. Paper prepared for the Committee on Programmed Instruction of the National Council of Teachers of Mathematics, 1964.

Holland, J. G. Evaluating teaching machines and programs. Teachers College Record, October, 1961, 63, 56-65.

Hovland, C. I.; Lumsdaine, A. A.; and Sheffield, F. D. *Experiments on mass communication*. Princeton, New Jersey: Princeton University Press, 1949.

Jacobs, P. I. Some relationships between testing and auto-instructional programming. *AV Communication Review*, November-December, 1962, *10*, 317-27.

Jacobs, P. I.; Maier, M. H.; and Stolurow, L. M. *A guide to evaluating self-instructional programs*. Princeton, New Jersey: Educational Testing Service, April, 1964. (Revised draft)

Klaus, D. J. The art of auto-instructional programming. *AV Communication Review*, March-April, 1961, *9*, 130-42.

Lindvall, C. M. (ed.) *Defining educational objectives*. Pittsburgh: University of Pittsburgh Press, 1964.

Los Angeles County Board of Education. An approach to the evaluation of instructional materials. *Our American Heritage*. Los Angeles: Our American Heritage Cooperative Project, May, 1963.

Lumsdaine, A. A. Partial and more complete automation of teaching in group and individual learning situations. *Automatic teaching: The state of the art*. (Edited by E. Galanter.) New York: John Wiley & Sons, 1959, pp. 147-66.

Lumsdaine, A. A. Some conclusions concerning student response and a science of instruction. *Student response in programmed instruction*. (Edited by A. A. Lumsdaine.) Washington, D. C.: National Academy of Sciences--National Research Council, 1961, pp. 471-527.

Lumsdaine, A. A. Criteria for evaluating self-instructional programs. *School Life*, November-December, 1962a, *45*, 15-17 ff.

Lumsdaine, A. A. Improving the quality of instruction through programmed learning. *Proceedings of the twenty-sixth conference of the educational records bureau and the American council on education*. New York: Educational Records Bureau, 1962b.

Lumsdaine, A. A. Report of the committee on auto-instruction programs and devices. *1962 annual report of the American Educational Research Association*. Atlantic City, New Jersey: American Educational Research Association, 1962c.

Lumsdaine, A. A. Some critical issues in the improvement of instruction through programed learning. *AV Communication Review*, January-February, 1962d, *10*, 61-64.

Lumsdaine, A. A. Some theoretical and practical problems in programmed instruction. *Programmed learning and computer-based instruction*. (Edited by J. E. Coulson.) New York: John Wiley & Sons, 1962e, pp. 134-51.

Lumsdaine, A. A. Assessment and improvement of instructional programs. *Educator*, December, 1963a, *6*, 4-6.

Lumsdaine, A. A. Instruments and media of instruction. Handbook of research on teaching. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963b.

Lumsdaine, A. A. Report of the committee on auto-instructional programs and devices. 1963 annual report of the American Educational Research Association. Chicago: American Educational Research Association, Spring, 1963c.

Lumsdaine, A. A. Some problems in assessing instructional programs. Prospectives in programing. (Edited by R. T. Filep.) New York: Macmillan Company, 1963d, pp. 228-62.

Lumsdaine, A. A. Educational technology, programed learning, and instructional science. Theories of learning and instruction: Sixty-third yearbook, national society for the study of education. Chicago: University of Chicago Press, 1964, pp. 371-401.

Lumsdaine, A. A., and Glaser R. (eds.) Teaching machines and programmed learning. Washington, D. C.: National Education Association, 1960.

Lumsdaine, A. A., and May, M. A. Mass communication and educational media. Annual review of psychology. (Edited by P. R. Farnsworth, Olga McNemar, and Q. McNemar.) Palo Alto, California: Annual Reviews, 1964, pp. 475-534.

Mager, R. F. Preparing objectives for programed instruction. San Francisco: Fearon Publishers, 1962.

Maier, M. H.; Stolurow, L. M.; and Jacobs, P. I. The schoolman's guide to programed instruction. Princeton, New Jersey: Educational Testing Service, February, 1963. (Mimeo)

Markle, Susan M. Confusion worse confounded. Contemporary Psychology, April, 1961, 6, 133-36.

Markle, Susan M. Good frames and bad: A grammar of frame writing. New York: John Wiley & Sons, 1964.

Miller, R. B. Handbook of training and training equipment design. WADC Technical Report No. 53-136. Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, June, 1953.

Miller, R. B. Task description and analysis. Psychological principles in system development. (Edited by R. M. Gagné.) New York: Holt, Rinehart and Winston, 1962, pp. 187-230.

National Education Association, Department of Audiovisual Instruction. Selection and use of programed materials: A handbook for teachers. Washington, D. C.: National Education Association, 1964.

National Society for Programmed Instruction. Checklist for selecting programs. NSPI Newsletter, October, 1962, 1, 6-7.

National Society for Programmed Instruction. Checklist for selecting programs. NSPI Journal, July, 1963, 2, 4.

New York City Board of Education, Division of Curriculum Development, Bureau of Audio-Visual Instruction. Recommended procedures for the appraisal of programed audio-visual materials. Brooklyn, New York: Programed Audio-Visual Materials Evaluation Project, 1962.

Paulson, C. F. The effectiveness of the self-instructional program longitude and latitude. Technical Report No. 1. Monmouth: Oregon State System of Higher Education, 1963.

Rocky Mountain School Study Council, Curriculum and Instruction Committee, Programmed Study Group. Guidelines for the selection of programmed materials, 1962.

Rothkopf, E. Z. Criteria for the acceptance of self-instructional programs. Improving the efficiency and quality of learning. (Edited by A. E. Traxler.) Washington, D. C.: Educational Records Bureau and American Council on Education, 1961, pp. 30-38.

Rothkopf, E. Z. Some observations on predicting instructional effectiveness by simple inspection. Journal of Programmed Instruction, Summer, 1963, 2, 19-20.

Rothkopf, E. Z. The epidemiology of ignorance. Contemporary Psychology, May, 1964, 9, 234-36.

Ryans, D. G. Report of the committee on auto-instructional programs and devices. 1961 annual report of the American educational research association. Chicago: American Educational Research Association, 1961, pp. 23-24.

Saltzman, I. J. Road under construction. Contemporary Psychology, March, 1961, 6, 106-109.

Schramm, W: Programed instruction. New York: Fund for the Advancement of Education, 1962.

Schramm, W. (ed.) Four case studies of programed instruction. New York: Fund for the Advancement of Education, 1964a.

Schramm, W. The research on programed instruction: An annotated bibliography. Washington, D.C.: U. S. Department of Health, Education, and Welfare, Office of Education, 1964b.

Schutz, R. E.; Baker, R. L.; and Gerlach, V. S. Teaching capitalization with a programed text. AV Communication Review, November-December, 1962, 10, 359-63.

Schutz, R. E.; Baker, R. L.; and Gerlach, V. S. Measurement procedures in programmed instruction. Tempe: Classroom Learning Laboratory, Arizona State University, 1964.

Shettel, H. H. Objectives, "ours' and theirs." NSPI Journal, May, 1964, 3, 12-14.

Silberman, H. F. Adventure or fundamentals. Contemporary Psychology, November, 1961, 6, 411-13.

Silberman, H. F., and others. Use of exploratory research and individual tutoring techniques for the development of programming methods and theory. Santa Monica, California: System Development Corporation, June, 1964.

Silverman, R. E. The evaluation of programmed instruction: A problem in decision making. Psychology in the Schools, January, 1964, 1, 74-78.

Skinner, B. F. Teaching machines. Science, October, 1958, 128, 969-77. [Reprinted in Teaching machines and programmed learning. (Edited by A. A. Lumsdaine and R. Glaser.) Washington, D. C.: National Education Association, 1960.]

Stolurow, L. M. Problems, procedures, pitfalls, and promise of programing practices. Bulletin of the National Association of Secondary School Principals, April, 1964, 48, 256-70.

Stolurow, L. M., and Lumsdaine, A. A. Training characteristics for illumination of animated classroom trainers. Technical Memo TARL-TM-56-1. Chanute Air Force Base, Illinois: Air Force Personnel and Training Research Center, Air Research and Development Command, Training Aids Research Laboratory, 1956.

Teaching Materials Corporation. How teachers can evaluate teaching machines and learning programs. New York: the Corporation, 1962.

Tracey, W. R. Editor's checklist: Content and construction review. NSPI Journal, December, 1963, 2, 13-14.

Tyler, R. W. Basic principles of curriculum and instruction: Syllabus for education 305. Chicago: University of Chicago Press, 1950.

University of Michigan, Center for Programed Learning for Business. Criteria for the evaluation of "off-the-shelf" programs. Programmed Learning Bulletin, April, 1963, 1, 1-2.

U. S. Air Force, Headquarters Air Training Command, Randolph Air Force Base, Texas. Evaluation of programmed instruction packages. Letter dated October 12, 1962.

Wiley, John & Sons. A guide for Wiley authors in the preparation of linear auto-instructional programs. New York: the Author, 1961.

Zuckerman, J. V., and Jacobs, M. Specific effects of the film "flight capabilities of the F86A" as related to points covered in the film. HRRL Memo Report No. 10. Washington, D.C.: U. S. Air Force Human Resources Laboratories, 1951.

CHAPTER IV COMPONENTS OF A SYSTEM OF INITIAL READING INSTRUCTION

John F. Cawley

The challenge to consider a program of Initial Reading Instruction is as exciting as the challenge is awesome. Because of the awesome nature of the task, the challenge cannot focus solely on reading. Initial Reading Instruction cannot be studied in isolation; it must be viewed as an integral component of a comprehensive system of learning. Such a system of learning must adequately provide for those children who demonstrate the behavioral prerequisites for satisfying and satisfactory participation in selected programs within the system. At the same time, the system must adequately provide quality experiences to make ready for participation those youngsters whose performance indicates they manifest developmental inadequacies.

A comprehensive system of learning should ultimately include (1) a hierarchial structure of skills and content areas that will assure ease of learning and effective transfer of learning within the system, (2) a continuous analysis of progress within the programs, (3) a constant feedback to the child and to the teacher of the results of his efforts and her programming, and (4) numerous program options suited to the various process strengths and weaknesses of the children, which the diagnostically oriented teacher may employ.

Originally prepared as a paper during the conduct of research on Reading and Psychomotor Disability Among Mentally Handicapped and Average Children, under a grant from the Connecticut Research Commission, by J. F. Cawley and H. A. Goodstein.

An effective system of learning will challenge the chronological basis upon which the curriculum of our schools and the expectancies of professional and community leaders are based. It shall accomplish this by making the individual the foundation of program development. It will clearly accept the fact that children are different, that they perform at different rates, and that they attain varying levels of accomplishment in various program areas. In effect, a comprehensive system of learning will accentuate differences, not homogenize them.

An effective system of learning will be of special significance for that segment of the school population classified as disadvantaged. Envisioned as beginning in the preschool years, it will bring structure and organization into the learning environment of the children. At the same time, such a system would initiate the process of diagnosis. This form of evaluation would identify specific behavioral deficits and prepare the way for programs of intervention and correction. The system would also provide for careful acceleration of the more gifted child.

To efficiently evaluate the place of initial reading instruction in a system of learning requires knowledge of specific perceptual, language and cognitive abilities underlying the beginning reading process. In addition, we should know the hierarchy of these abilities and the relationships between maturation and learning that maintain in this area.

The ability of the child to use reading as a meaningful source of information and to "comprehend" must depend, in addition to specific perceptual skills, upon levels of language development and general cognitive abilities. Many children possess mature language structure at the age of four, only to await minor refinement of syntax and vocabulary growth (Berko, 1958). However, we do know that here exists wide individual differences in the timing of acquisition of mature language structure, while some children, because of their language environments, never attain standard English.

The relationship between general cognitive abilities and success in early reading has been extensively researched. In general, moderately high correlations between mental age and subsequent reading achievement, resembling those correlations for perceptual discrimination tasks, have been found (Allen, R.J., _et al._, 1959; Dean, 1938; O'Sheasey, 1951;

Pratt, 1949). However, when an intelligence test has been included with a battery of tests measuring various aspects of visual and auditory perception, most of the variance accounted for by the intelligence test is also accounted for by the various perceptual tests (Barrett, 1965b; Harrington and Durrell, 1955; Silberberg, Levenson and Silberberg, 1967). Two interpretations may arise for consideration. First, the construction of first grade or kindergarten level intelligence tests or test items may stress those very areas of perceptual skill tapped by other measures. Second, it may be that a great deal of initial reading success is determined by perceptual readiness, while cognitive abilities play an increasingly important role in predicting reading achievement at later stages of reading when comprehension is important.

A mass of research on the relationship of reading readiness to initial and later reading achievement exists in the literature (Bagford, 1968; Robinson and Hall, 1942). The emphasis in this research has been toward looking at the correlation of the total score on the test subsequent achievement. The desirable goal is to improve the placement of children in beginning reading programs. Very often, however, correlations with reading achievement are too low for effective prediction.

The role of memory, short-term retention, associative learning and numerous other factors in initial reading instruction is far from clear. The nature of the integration of the auditory/visual processes and the relationship of these to each of the other previously mentioned components (behaviors) is in need of further clarification. More comprehensive and definitive factorial studies and multivariate strategies with the components of a system of learning will gradually clarify some of these questions. The complexity of the relationships among these processes is stressed in the following statement by Wesman (1968):

> If we accept the thesis that the modules or bits which constitute intelligence are themselves complex, a combination of such modules can hardly be expected to be simple or pure.
>
> A 6-year-old who assembles three alphabet blocks to spell out "cat" has employed, at a minimum, verbal and spatial skills; if he is aware that there are three blocks or letters, he has engaged in numerical perception as well. The ability to perform the task has

> required cognition, memory, convergent thinking, and evaluation. The product is figural, symbolic and semantic. All this, and we have not yet taken in account such considerations as motor-manipulative activity, the perception of color, the earlier learning experiences which enabled him to perform the task successfully, or the imagery which the concept "cat" induces in him.
>
> We, as analysts, may choose to attend only to a single aspect of behavior but the behavior itself remains multifaceted and complex.

Wesman's conceptualization of intelligent behavior must be kept in mind as educators move toward a more involved analysis of initial reading instruction and expand efforts to prevent failure.

The extent to which we can train perceptual abilities in the hierarchy of tasks in a system leading to learning to read is an important issue in education today. To the degree that we can structure the learning situation for the child to speed up the maturation of these functions and refine and direct their development, preschool and kindergarten education will take on the look of a diagnostic-prescriptive learning situation. In addition to the general question of training perceptual abilities must be added the question of the type of training most appropriate and efficient. For example, should we employ traditional discrimination training (i.e., matching to a standard) or original learning activities (i.e., two choice discrimination learning tasks)? At what level is recognition training more facilitating upon performance than discrimination training?

One ability prerequisite to efficient learning to read is visual discrimination. Skilled readers do not rely solely on the perception and discrimination of individual letters in order to recognize words (Neisser, 1967). However, much is yet to be learned of the organizing principles of initial learning to recognize words. One important aspect of this is visual discrimination of letter shapes. The ability to discriminate letters and, thus, words has been shown to be a highly significant predicator of first grade reading achievement (Barrett, 1965a; Shea, 1968).

Barrett (1965a), in an exhaustive review of the literature, found the ability to discriminate verbal materials (letters and words) in general to be a better predictor of reading achievement than ability to discriminate non-verbal materials, although relationships between the latter and first grade reading achievement are by no means absent. Goins (1958) and de Hirsch (1966) found moderately big correlations between even more primitive levels of visual perceptual organization and reading achievement.

If one views perceptual learning as stratified, proceeding from gross to finer levels of functioning, the data lend themselves to a reasonable analysis. Visual discrimination of letters and words, being closer to criterion abilities directly underlying initial reading instruction, will correlate higher with beginning reading than an ability to perform a visual perceptual task at a more gross level, simply because the former ability requires attainment in the prerequisite task. Marchbank and Lewis (1965) concluded that the first letter and the last letter in the word are the primary cues used by non-readers and by beginning readers to remember a word; shape is the least used cue.

Proceeding with this analysis, it comes as no surprise that Barrett (1965b) and Silberberg, Levenson and Silberberg (1967) report that by far the strongest predictor of initial reading in their respective batteries was the ability to name letters. This performance requires the prerequisite performance of learning of visual discrimination of letter shapes with the additional association of shape with a letter name. To the extent that this ability underlies success in word recognition and reading, the high relationship is evident.

Samuels and Jeffreys (1966) studied the discriminability of words and letters in learning to read. The investigators note that reading texts select words relative to their frequency of usage. This increases the likelihood of the child being familiar with the sound of a given word. It is pointed out that high frequency words tend to be dissimilar and that this dissimilarity should aid in initial learning. However, initial reading instruction should enable the child to attain a high level of transfer from the initial training to new and unfamiliar combinations. It is questionable that they do. In order to study this, thirty-six kindergarten children were confronted with a paired associate task in which they were to learn three lists of words that had been constructed from an artificial alphabet. One list had four two-letter words constructed from four letters; the second list had four two-letter words constructed from six letters; and the third list had four two-letter words constructed from eight letters. The data indicate that the number of children who make identifications on the basis of a single letter increases with the number of letters on which they were trained. Training that focuses attention to each letter is less likely to lead to subsequent reading errors than training which permits a child to identify words on the basis of a single letter. It is suggested that children have experience identifying similar words which forces identification on the basis of more than a single stimulus feature.

Tachistoscopic training of the recognition of capital letters, an association task, has been found to improve future performance by kindergarten children on a multiple-choice matching visual discrimination task with letters. In this study (Wheelock and Silvaroli, 1967), the performance of children from lower socioeconomic classes was especially enhanced. Popp (1967) has demonstrated that a program of multiple-choice matching tasks involving reversible letters, bigrams and trigrams can significantly improve visual discrimination ability in an experimental group of kindergarten children. Popp also notes that the correlation of discrimination tests scores with later reading achievement may indicate that an ability to discriminate does influence reading achievement or that some underlying common factor exists which produces high scores on both measures; the same might be said in the case of low scores. Effective programming that provides a means of observing and controlling a subject's interaction with specific instructional materials will assist in a greater understanding of these issues.

Goins (1958) found practice on tachistoscopic perception of numbers could improve the span of apprehension for numbers in an experimental group of children in the first grade; however, no significant improvement in reading achievement was found for the experimental sample.

Since reading involves the decoding of visual stimuli into auditory language patterns, auditory perceptual abilities should also be related to efficient learning to read. Auditory discrimination of words has been found to be a moderately high predictor of initial reading achievement (Hanesian, 1966; Harrington and Durrell, 1955; Nila, 1953; Thompson, 1963;

Wepman, 1960). However, little is known about the relationship of auditory abilities, at the more gross levels,with subsequent auditory discrimination of words and ultimately reading achievement.

The development and role of auditory perceptual skills in reading is another area of considerable importance at the level of initial reading instruction.

McNeil and Stone (1965) trained kindergarten children on nonsense syllable and regular words to distinguish sounds in spoken words. Those trained on nonsense syllables produced fewer errors and more correct responses than those trained on regular words. It is possible that they were distracted by the meaning in the regular words and, therefore, did not attend as well to specific sounds.

They are important because there may be a time during which the auditory skills are maximally related to reading skills (Feldman and Deutsch, 1966) and a deficit at that point might effect subsequent reading skills. Precisely when this point is reached is uncertain. Katz and Deutsch (1964) showed a greater differentiation between good and poor readers at the first grade level than at the third grade level on auditory skills. Bryan (1964) found a much greater relationship between visual perceptual skills and reading at the beginning levels than at the upper levels. By third grade, intelligence appeared to relate more significantly to reading. It seems possible that some children obtain an optimal age at which a basic skill such as auditory discrimination generalizes automatically or effectively to reading, and because of this and the fact that training in auditory perceptual skills does not seem to effect the reading achievement of third grade children. Feldman and Deutsch (1966) suggest that training in the perceptual skills should precede reading training. The potential for prevention in initial reading instruction is obvious.

One study of auditorv perceptual training (Feldman and Deutsch, 1966) assessed the impact of this training on children. In this study, Puerto Rican children had significant interaction with auditory play and Negro children had positive interaction with the reading play treatment. Puerto Rican children, who received auditory training only, performed better on reading than Negro children who received auditory training only. Negro children, on the other hand, who received reading instruction, responded better to reading than Puerto Rican children who received reading instruction. This certainly highlights the need for systems of learning which contain programs suitable to specific children.

Monroe and Rogers (1964) stress the importance of oral language and listening skills in the pre-reading program. Differences in the characteristics of sounds is a sensitivity which these writers feel children should develop. Consideration is given to such topics as Intensity, Pitch, Quality and Duration and Sequence. To illustrate, the child's attention to intensity might be developed by hiding an object and having the child seek the object with the aid of cues from the class. As he approaches the object, the class can clap loudly and as he moves away, the clapping would be softer.

In regard to sensitivity to sounds in words, Monroe and Rogers approach the delicate topic of phonetic drills by suggesting that teachers often drill youngsters at levels beyond those at which the youngsters can make auditory discriminations and associations. Factors such as rhyming words, alliteration and responsiveness to isolated speech sounds should be adequately developed in each child. A major problem is that we lack a clear understanding of the contribution of the auditory processes to reading and, furthermore, we lack a clear description of the hierarchy of auditory prerequisites required to make a child skillful at any given level.

Curriculum development in auditory discrimination has received only minor attention. Feldman and Duetsch (1966) developed an auditory perceptual training program for use with disadvantaged children. This curriculum included sound recognition, sound discrimination, auditory memory and attentivity. In this program the same auditory skills were taught in the same sequence by all tutors in the study. Among the activities included were (1) <u>environmental sounds</u>; identification of environmental sounds; (2) <u>following directions</u>; the child was given oral directions and he carried out an assigned task; (3) <u>words</u>; this included the repetition of words and rhymes; (4) <u>sounds of letters and letter combinations</u>; child supplied words which had given sound or they learned to associate letter sounds and names; (5) <u>blending sounds</u>; child blended sounds without the aid of visual cues; (6) <u>listening to stories</u>; and (7) <u>telling stories</u>.

A basic program in auditory discrimination has been developed at the University of Connecticut (Kummer, Smith, Carrington and Bilodeau, 1968). The strategy used in developing this program was to (1) identify selected components of auditory discrimination, (2) review the experimental literature, and (3) plot a sequence of discrimination activities. The curriculum did not originate with the idea that a particular number of lessons were to be devoted to a particular topic. Rather, the procedure was to plot the number of lessons judged necessary to develop a defined behavior and to prepare a program that would develop this behavior. The end result was a program consisting of five units and ninety lessons. The units are (1) Gross Sounds, which had five units and a total of fifty-five lessons, (2) Spondee Words, with five lessons in one unit, (3) Nonsense Syllables, grouped in five units and twenty lessons, and (4) two units which focussed upon meaningful words and a higher level of nonsense syllable usage. The important aspects of this program are the facts that it was developed by teachers as a meaningful curriculum entity and the stress placed upon a preventive orientation.

Bateman's (1967) research on modality effectiveness and differential programming with first grade children indicates that the auditory oriented programs are substantially more beneficial than visually oriented programs. Bateman identified children whose scores on the auditory sequencing and visual motor sequencing memory tests of the ITPA indicated a modality strength in either the visual or auditory processes. The overall auditory abilities of the youngsters was approximately nine months higher than their visual abilities. Those youngsters who scored nine or more months higher on auditory tests than in visual were classified as auditory modal and reading instruction was provided through an auditory program. Those youngsters whose auditory memory was lower than nine months above visual memory were classed as visual modal and were taught reading through a more visual approach. Two other samples of mixed subjects were taught with one of the two approaches. Auditorally modal subjects scored higher than visually modal subjects in their respective programs and the auditory methods seemed superior to the visual methods in the mixed groups.

A valuable addition could have been made to the study had the relative strengths of the two modalities contrasted in the four groups been equated.

The effect of readiness programs upon subsequent achievement in beginning reading cannot be totally evaluated because of the inherent structure of these programs. Most programs have attempted to introduce training in various areas of "readiness." Evaluation of the program is made employing univariate means (i.e., program versus no program, structured versus unstructured) in an area where the problem investigated, readiness, is multivariate. Future programs must include a means by which one could check the progress of the children in various aspects of the program. Future research with these programs must employ these measures in their analysis. Otherwise, the issue as to which aspect of the program, which combination of abilities, or whether the total program had an effect upon subsequent achievement is not dealt with.

As with the specific training programs of perceptual abilities, general readiness programs appear less effective when employed in conjunction with the regular first grade reading program (Niles, 1966). This study was performed with children who were potential problem readers. However, as with the Jordan study positive results from this training may appear at later stages in the reading program.

We now have an extensive body of literature comparing different methods and materials used in initial reading instruction in the first grade (Bliesmer and Yarborough, 1965; Bond and Dykstra, 1967; Chall, 1967; Gurren and Hughes, 1965; Woodcock, 1967). The results are often contradictory and confusing. In general, however, it may be concluded that early instruction in phonic skills leads to significant improvement in word recognition skills. The effect of various programs on reading comprehension supports no reasonable conclusion at this time.

Studies done comparing approaches to initial reading instruction are beset by methodological problems. In comparing methods most often different materials are used. The effect of methods and materials upon subsequent achievement is necessarily blurred. Many reading programs emphasize different components of the reading process according to different time tables; often first grade achievement scores give a distorted picture of the total outcome of a reading method. Extension of many of the first grade reading studies sponsored by the U. S. Office of Education into longitudinal studies appears a wise decision. Studies contrasting use of an artificial orthography (I.T.A.) with use of a traditional orthography, appear especially handicapped in the early years of evaluation, when evaluation is made with achievement tests employing traditional orthography (Woodcock, 1967).

The data from those first grade reading studies (Reading Teacher, May, 1966; October, 1966; and May, 1967) suggest that some children, a potentially identifiable proportion within each sample, did not demonstrate maximum responsiveness to any program. It could be suggested, therefore, that the differences between one approach and another are not nearly as important as the differences between the responsive and unresponsive children in each treatment. The very design of these studies, evaluated in terms of mean differences in reading achievement, may not be appropriate for evaluating the most advantageous approach to initial reading instruction. This design, which pits method A against method B, irregardless of the profile of abilities of the children involved, especially for the child of low abilities, inherently works against the teacher individualizing her program. The number of program options are limited by the necessity of keeping method A separate from method B. Here again, we see the inadequacy of approaching multivariate problems with univariate solutions.

The philosophy behind the embarkment upon this course of research is that there is one method that will surely prove itself superior with all children. Very often this is translated to mean one program will prove most advantageous with all children. This philosophy may be challenged on the grounds that children are different, they enter the reading situation with vastly different profiles of abilities, and may profit from one type of learning situation much more than another. The child of average or high ability will probably be able to benefit from any reasonable program offered him. The child who is likely to encounter difficulty in reading will probably need more options for learning than is offered in any one program. If he didn't, he wouldn't be failing in reading. Yet the exploration of the types of options required for different children of differing ability patterns awaits to be explored. Blind commitments on the part of reading researchers to the belief that one program or approach will eventually be found to be the panacea for the teaching of reading will further retard meaningful research into initial reading instruction.

How does one initiate a beginning reading program for disadvantaged children? Bereiter and Englemann (1966) posit that reading instruction should familiarize the child with the alphabetical principles. This necessitates teaching him that words are spelled. According to this notion, if the child is to make progress, he must clearly recognize that it is spelling and not some superficial characteristic like length or symmetry that creates a word. Readiness should include the development of an awareness of words as distinct entities and an awareness of the alphabetical principles; this requires that the child be taught the letters. This program has two main components. One of these is the development of an awareness of words, word recognition and the eventual reading of sentences. Contiguous with this is letter discrimination and identification.

Bateman (1967) wants reading taught as a rote, conditioned, mechanical process of converting letters to sound. Comprehension would be taught as a separate process.

Engelmann (1967) suggests that disadvantaged children with mental ages of four years and above can successfully be introduced to formal reading programs. Fundamental to this type of programming are teaching the child (1) that the spoken word is composed of parts, (2) that the parts occur in fixed order in time, and (3) that the reading code represents that passage of time through a left to right progression of symbols. The unique aspect of the program proposed by Engelmann is not in methods, but in structure and in the stress placed upon the acquisition of skills which have relevance to reading and the willingness to break these skills down into subskills that will yield a given product.

Readiness is not easily defined. As used in this context, readiness is contiguous with readying the child to learn. This infers that programming and instruction are two elements which enhance the child's status for successful performance of a task. Within this frame of reference, an essential component of readiness is diagnosis. Diagnosis requires that the instructor understand and be alert to the stimulus properties and performance requirements for each task with which the child is confronted. This is probably the big advantage of phonics in contrast to "look and say;" basals are often so gross that stress on needed skills are masked.

The child's performance must be capable of examination and recording and adaptation of the program must be as constant as diagnosis indicates.

No doubt the reader senses a tendency to reject the notion of developmental-readiness. The tendency to reject this construct is based upon the proposition that programs must contain quality experiences for all children. We are in agreement with the writing of Engelmann (1967) where he suggests that when a certain number of children do not perform well, there is a tendency to excuse the program and focus on the deficiencies of the children. Accordingly, these children can be viewed in two ways; as children who fail because they have not received adequate instruction, or as children who fail because they lack aptitude, readiness or intelligence. Failure due to lack of aptitude and readiness can be explained by the developers of the program in the sense that their program did not provide for these children. In neither instance was the program responsible or inadequate. The question then must be asked, "Why were they in it?" If programs were constructed to provide a diagnostic base for entrance, there might be fewer instances of this kind. At the same time, programs must provide enough workable alternatives to enable the teacher to divert the youngster into other programs. This supports the notion that systems of learning must be developed and the notion that composition of our reading programs is entirely inadequate in this day and age.

One of the most serious obstacles confronting the diagnostically oriented teacher in kindergarten or first grade is the lack of continuous observations and records of the child's performance. A substantial portion of commercial materials fail to control the stimulus properties of the materials, fail to develop stability of performance and are inadequately organized to permit diagnosis and intervention. Without early intervention, the preventive component of the program is diminished and children are going to be wallowing in failure.

The programmatic concepts specific to academic readiness are undergoing extraordinary transition. The advent of Project Head Start suggests that diagnosis can be introduced earlier in the education of disadvantaged children and the steps to amelioration of specific and generalized deficits undertaken. In fact, de Hirsch (1966) goes so far as to suggest that maturation is contingent upon functioning which is fostered by experience and training. The additional year of education - year four - provides for earlier diagnosis, intervention and initiation of a continuum of educational experiences specific to the characteristics of the child. Head Start possesses the capability to provide the kindergarten with a concise description of each child with the possibility that an immediate continuance of psychoeducational teaching will be maintained. This contrasts with previous years when the kindergarten teacher was a substantial part of the way through the year before she was able to detect deficiencies with a child. The transmission of the form of programming to the first grade teacher will make her ability to adapt programs to each child easier. This will also create problems for the primary teachers because if large numbers of children begin to derive sustained benefits from their preprimary experiences, the class differences will be accentuated and the problem of programming will be more difficult.

The position that the child is inadequate for participation in the initial reading instruction is a false generalization. Readiness which is task specific will require that we ask ourselves, "What is he ready for?" "What are his specific areas of performance strength and weakness?"

We find it difficult to predict success and failure in reading. As a result, the preventive components of an initial reading instruction are hampered.

Components of an Initial Reading System

This section will focus on four components of an initial reading system. They are (1) hierarchical sequences of skill development, (2) diagnostic teaching, (3) alternatives for individualization, and (4) corrective intervention.

Hierarchical sequences of skill development developed within the framework of a structured program is an essential component of an effective system of learning. Various strategies for training specific tasks must be evaluated as to the prerequisite abilities required for maximum probability of success for the child in developing each new ability. Further, the task must be broken down into as many steps as is necessary to complete the job without leaving "soft spots" in the system that will retard the child's later performance in a subsequent learning situation. The tasks should be sequenced in order to maximize, at the beginning, the child's probability of success. Perhaps discrimination tasks should begin as two-choice problems, then gradually increase the number of choices and the similarity between the choices for the child.

Provision must be made within the system for diagnostic teaching. It is encumbent upon the developers of initial reading programs to ascertain the degree to which incipient and manifest disabling conditions can be properly identified. The substantial majority of primary grade reading programs are developmentally based within the framework of chronological age. Youngsters are incorporated into the system via a direct confrontation with generalized curriculum and methodological strategies. Diagnosis generally follows some deviation from the expected pattern. While it is argued that the shortage of funds and personnel are major obstacles in the development of pre-failure diagnostic programming, there is a paucity of this type of programming built into most of the current reading programs or into the teacher training programs. Effective pre-failure diagnosis would lead to proper planning for the child. Proper planning could avoid much of the failure experience often associated with initial reading instruction. Prevention would usurp remediation.

Preschool and kindergarten programs should include a comprehensive diagnostic component. The tasks which are planned should provide for continuous assessment and diagnosis to identify performance deficits in a variety of behavioral areas. The training of teachers of young children should focus more on psycho-educational strategies than methods. Methods used with each child will only be as effective as the validity of the psycho-educational strategy underlying use of that method with each child.

A system of initial reading instruction by its very organization provides alternatives for individualization of instruction. The ultimate in educational planning will not become a reality until the system of learning provides these alternatives. This may require architectural innovations which will permit the simultaneous deployment of children into multiple learning activities. In effect education might take place in resource centers rather than self-contained classrooms. Since the system will require numerous options in promoting the desired abilities, attention will have to be placed initially upon incorporating the child into the system and teaching him how to function within it. This might involve teaching the child how to play the learning game, described as the development of learning set in the experimental literature. Computer-Assisted-Instruction, auto-educational materials, films, filmstrips, slides, tapes, and other forms of instructional materials could be made an integral part of the system.

As the system progresses a greater responsibility will rest on the teacher. She will be responsible for guiding the youngster to the proper programs, for assessing the impact of these programs upon the child, and for the channeling of the child into a proper psycho-educational learning style. The question of whether a child might learn to read via an analytic or synthetic approach to phonics or some other selected tactic will be determined by the teacher based upon prior performance patterns and continuous diagnostic teaching. Research into some of the outcomes of this approach in beginning reading instruction, although utilizing only one criterion for placement, has already begun (Bateman, 1967).

Corrective intervention, as conceived in this system, relates to the ability of the teacher to re-channel performance programs for the child to enhance a greater proportion of success experiences. We would not be dealing with long-range prediction, but with day-to-day analysis and program adjustments. This, hopefully, would prevent failure from becoming the predominant pupil-to-teacher-to-stimuli mode of interaction.

Corrective intervention does not infer "catching up." What is intended is that the program will alleviate developmental deficiencies of the child at the level at which he is functioning. If this clears the way for rapid progress, then this is proper. The question must not be whether an eleven year-old child reads at the third grade level. It must be whether or not he is reading at that level because the program in which he is participating successfully brought him to that level in a manner that made his learning efficient, effective, and enjoyable. The sorrowful aspect of having a group of children reading at one level, while their peers read at a higher level is the fact that the curriculum and program anticipated that they would be reading at the higher level and the program failed to get them there. In an effective initial reading system, the child would be introduced to successful learning and would derive satisfaction from this regardless of his reading level, not dissatisfaction as a result of failure, repitition and chaos of a program that was not designed for his being at that level.

The initial reading system is most necessary for the disadvantaged child. An effective start in reading will intervene with the cumulative deficit pattern often associated with these children (Deutsch, 1966). It is argued (Bereiter and Engelmann, 1966; Ellis, 1966) that an academic orientation to preschool is a necessity if the disadvantaged child is to be competitive upon entrance into school. Bereiter and Engelmann suggest a comprehensive approach to the efficient use of time spent in preschool by selecting experiences that produce more learning. The system would provide a means for teachers of kindergarten and the primary grades to follow up learning patterns developed from the preschool.

Cawley Early Performance Program

The Cawley Early Performance Program (CEPP) is being studied as one possible component of the system of learning that has been referred to in this paper.

CEPP is a diagnostic teaching program which is designed to provide a variety of opportunities for teachers and teacher assistants to observe and assess the performance of children. The program attempts to delineate specific tasks which can be sequentially and developmentally used with children from about forty-two mental age months to approximately seventy-eight mental age months. CEPP is predicated upon the notion that diagnosis and teaching are inextricably interwoven, particularly in programs for disadvantaged and handicapped children which mandate individualized instruction.

CEPP is not academically oriented in the formal sense. It is structured and it does focus upon the assessment and remediation of specific performance disabilities. Currently there are eighteen task categories which include approximately ninety different types of stimulus/response combinations. The projection is that the total program will consist of approximately six hundred lessons.

CEPP is not intended for use with any particular etiological or diagnostic sample of children. Rather, its sequential pattern suggests that it is appropriate for use with any youngster, or group of youngsters, who exhibit or may potentially exhibit performance disabilities. Some groups of children (e.g., the culturally disadvantaged, the mentally handicapped, the minimally brain-injured) manifest a greater concentration of performance disabilities than do those children who are similarly impaired.

Teacher information forms the basis for the tactic. With adequate information, the teacher is more effective in the diagnostic-teaching role. This enables her to observe the behavior, or the results of behavior, of the child and to simultaneously plot and conduct her diagnosis. Teacher information is also the basis for the selection of the tasks which the youngster is to be confronted with. The task, and the criterion established for each task, is a prelude to the performance of the child.

In CEPP teacher information forms the basis for the selection of an appropriate task for the youngster. This combination is the main constituent of the diagnosis, which is the function of the child's performance. For this reason, each task is as discrete as rational judgment permits.

The Tasks. Each task with which the youngster is confronted is described in two ways. The first is the task category, which is a broad heading for the type of behavior that will be elicited from the child (e.g., visual discrimination, auditory discrimination, visual memory). Each of the categories is subdivided into selected relevant tasks. One of these is auditory discrimination (Figure 1) in which the youngster is asked to mark the one that goes with the word that is presented by the teacher (e.g., "Mark the sock."). The complexity of the discriminability between and among items increases throughout the program. In contrast with the simple presentation in block 1 of Figure 1, block 5 increases to a pair of words (e.g., "Mark lock-sock."). The differences between blocks 2 and 3 and block 4 and between blocks 5 and 6 and block 7 are not clear from the figure. Descriptively, the stimulus pair is randomly arranged throughout the lesson, but they remain the same. This is true for block 3 also. In block 4, the four stimuli from the two lessons are randomly mixed. The criteria for each task varies with the number of options. Many of the activities have fifteen plates to a lesson. On two-choice options, the child is required to make thirteen correct choices; three-choice options require eleven correct choices; and four-choice options require nine correct choices. In any two-choice task, fifty per cent, or

SAMPLE ITEMS OF CEPP AUDITORY DISCRIMINATION

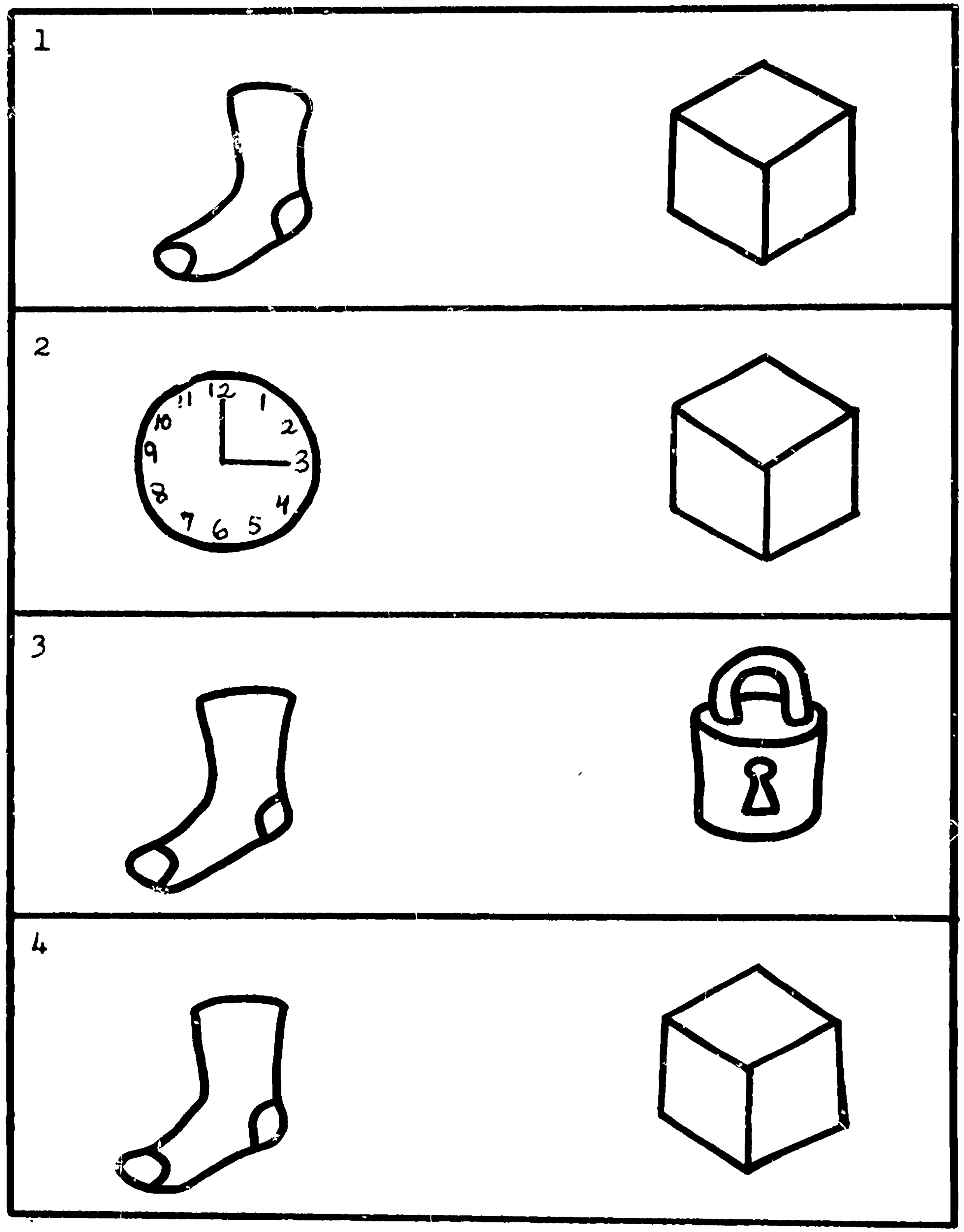
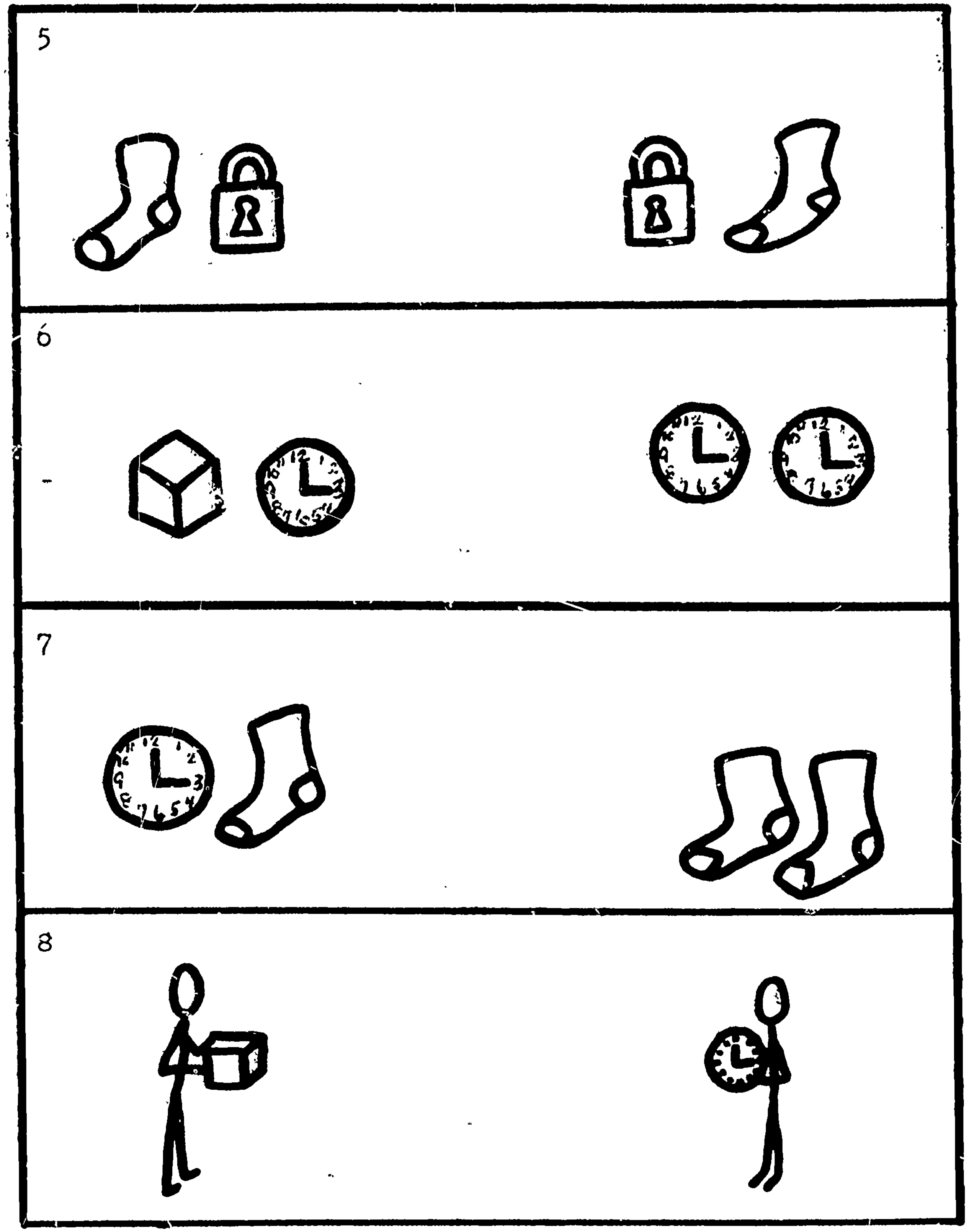


Figure 1 (Continued)



about eight in this program, would simply be chance. The program, by establishing reasonable criterion, encourages beyond chance response patterns. The young child, or the child with specific performance defects, is often inconsistent in his performance. The program attempts to compensate for this. Thus, blocks 2 and 3 work on the stability factor, whereas block 4 is a review task.

The instability in the response pattern of the young child is a process which should be observed. In order to approach this problem, in addition to establishing specific criterion, a substantial number of lessons in the early phase of the program maintain the same stimuli on each plate. As the program progresses, there is considerable variation from block to block on each task.

CEPP requires each child to make a written record of his performance. In some cases, this simply calls for the child to make a mark, while in others he might complete a maze or make a simple drawing. This output mode provides the teacher with a record of performance that she can examine and analyze. This seems preferable to those programs wherein a child holds up his hand or utilizes some other means of expression because of the difficulty the observer has in making a record of the stability of the child's performance.

Presenting the Program. The program can be used with individuals, small groups, or an entire class. In the formative stages, it is suggested that the maximum group size be three or four. Once the youngsters have grasped the idea of the program, selected lessons, and they are labeled, are appropriate for use with the class as a whole.

The teacher should be responsible for selecting the task for each youngster, although the teacher's assistant can administer the program and score and graph each day's performance. The teacher analyzes the graph and makes the decision as to the next step. CEPP has five or more lessons which focus on the same task. The purpose for this is to provide a basis for continuing a youngster in the same performance category, should it be advisable.

It takes about ten minutes for a youngster to complete most of the tasks in CEPP. As youngsters become more sophisticated in its use, they might go directly to the program during a lull and complete their assigned tasks.

CEPP's potential as a facilitator of performance will be enhanced if each lesson, or at least selected lessons, is concluded with a brief discussion of what took place. Some teachers will certainly want to conduct pre-participation discussions with the children on some of the lessons. It would be proper, in either case, to note the kinds of tasks in the lesson and to interpret these with the children. The task in Figure I might be presented as follows:

Teacher: Who can tell me what you were doing in today's game?

Child I: We were making marks on the pictures.

Teacher: On what pictures were you making marks?

Child 2: On the ones that go with the word that you said.

Teacher: That's right, we were listening for the differences between sounds (words).

Summary

In summary, the authors have attempted to review the clinical and experimental components of initial reading instruction. Research relative to these areas has been presented and discussed. Our present state of knowledge is such that a blind commitment to any particular strategy is unwarranted. These are identifia le correlates of initial reading. They are correlates and not causes. Research has yet to determine why one youngster possesses a set of characteristics that enables him to successfully learn to read. The research that has described the abilities and disabilities among good and poor readers has not contributed substantially to our knowledge beyond the descriptive stage.

One program, CEPP, which is currently being tested for possible inclusion in a comprehensive system of learning was discussed. This program is predicated upon the notion that if the correct parameters can be properly identified and sequenced, and children are able to respond to a program of about six hundred lessons, a successful basis for learning to read could be established.

References

Allen, R. J., et al. The relationship of readiness factors to January first grade reading achievement. Unpublished master's thesis, Boston University, 1959.

Bagford, J. Reading readiness scores and success in reading. Reading Teacher, 1968, 21, 324-328.

Barrett, T. C. The relationship between measures of pre-reading visual discrimination and first grade reading achievement: A review of the literature. Reading Research Quarterly, 1965, 1, 51-76.

Barrett, T. C. Visual discrimination tasks as predictors of first grade reading achievement. Reading Teacher, 1965, 18, 276-282.

Bateman, B. Reading: A controversial view, research and rationale. Curriculum Bulletin, 1967, 23, 1-14.

Bereiter, C. and Englemann, S. Teaching disadvantaged children in pre-school. New Jersey: Prentice-Hall, 1966.

Berko, J. The child's learning of English morphology. Word, 1958, 14, 150-177.

Bliesmer, E. P. and Yarborough, B. H. A comparison of ten different beginning reading programs in first grade. Phi Delta Kappan, 1965, 46, 500-504.

Bond, G. L. and Dykstra, R. Interpreting the first grade reading studies. In R. Stauffer (ed.) The first grade reading studies: findings of individual investigations. New York: IRA, 1967.

Bryan, Q. Relative importance of intelligence and visual perception in predicting reading achievement. California Journal of Educational Research, 1964, 15, 44-48.

Chall, J. Learning to read: The great debate. New York: McGraw-Hill, 1967.

Dean, C. D. Predicting first grade reading achievement. Elementary School Journal, 1938, 39, 609-615.

de Hirsch, K., Jansky, J., and Langford, W. Predicting reading failure: A preliminary study of reading and spelling disabilities in preschool children. New York: Harper and Row, 1966.

Deutsch, M. Some psychological aspects of learning in the disadvantaged. Teachers College Record, 1966, 67, 260-265.

Ellis, R. Educational programming for preschool children. Child Study, 1966, 28, 23-32.

Englemann, S. Classroom techniques: teaching reading to children with low mental ages. Education and Training of the Mentally Retarded, 1967, 2, 193-202.

Feldman, S. and Deutsch, C. P. A study of the effectiveness of training in retarded readers in the auditory perceptual skills underlying reading. New York: Institute for Developmental Studies. New York Medical College.

Goins, J. T. Visual perceptual abilities and early reading progress. Supplementary Educational Monographs, 1958, 87, 1-108.

Gurran, L. and Hughes, A. The Cleland-Miller study on phonics. Elementary School Journal, 1965, 66, 87-91.

Hanesian, H. The relationship of auditory abilities to first grade reading achievement. Unpublished master's thesis, Columbia University, 1966.

Harrington, Sister Mary and Durrell, D. Mental maturity versus perceptual abilities in primary reading. Journal of Educational Psychology, 1955, 46, 375-380.

Katz, P. and Duetsch, M. Modality of stimulus presentation in serial learning for retarded and normal readers. Perceptual and Motor Skills, 1964, 19, 227-233.

Kummer, H., Smith, S., Carrington, G., and Bilodeau, K. Auditory Discrimination. Storrs: University of Connecticut, 1968. (mimeo)

Marchbanks, G. and Lewis, H. Cues by which children recognize words. Journal of Educational Psychology, 1965, 56, 57-61.

McNeil, J. and Stone, J. Note on teaching children to hear separate sounds in spoken words. Journal of Educational Psychology, 1965, 56, 57-61.

Monroe, M. and Rogers, B. Foundations in reading instruction. Chicago: Scott-Foresman, 1964.

Neisser, U. Cognitive psychology. New York: Appleton Century Croft, 1967.

Nila, Sister Mary. Foundations of a successful reading program. Education, 1953, 73, 543-555.

Niles, O. Methods of teaching reading to first grade children likely to have difficulty in reading. Reading Teacher, 1967, 20, 541-545.

O'Sheasey, E. A. The predictive value of certain pre-reading measures employed at the first grade level. Unpublished master's thesis, Boston University, 1951.

Popp, H. The measurement and training of visual discrimination skills prior to reading instruction. Journal of Experimental Education, 1967, 35, 215-216.

Pratt, W. E. A Study of the differences in prediction of success in kindergarten and non-kindergarten children. Journal of Educational Research, 1949, 42, 526-533.

Robinson, F. R. and Hall, W. E. Concerning reading readiness tests. Bulletin of the Ohio Conference on Reading, Ohio University Press, 1942.

Samuels, S. and Jeffrey, W. Discriminability of words and letter cues used in learning to read. Journal of Educational Psychology, 1966, 56, 337-340.

Shea, C. A. Visual discrimination of words and reading readiness. Reading Teacher, 1968, 21, 361-368.

Silberberg, N., Iverson, I., and Silberberg, M. Predicting end of the first grade developmental reading test scores from a reading readiness test administered in kindergarten. Proceedings of the 75th Annual Convention of the American Psychological Association, 1967, 2, 291.

Thompson, B. B. A longitudinal study of auditory discrimination. Journal of Educational Research, 1963, 56, 376-378.

Wepman, J. M. Auditory discrimination and speech and reading. Elementary School Journal, 1960, 60, 325-333.

Wesman, A. Intelligence testing. American Psychologist, 1968, 23, 267-275.

Wheelock, W. H. and Silvaroli, N. Visual discrimination training in beginning readers. Reading Teacher, 1967, 21, 115-120.

Woodcock, R. The Peabody-Chicago-Detroit reading project--a report of the second year results. Institute of Mental Retardation and Intellectual Development, 1967, 4, 1-15.

CHAPTER V COMPUTER-ASSISTED INSTRUCTION

Max Jerman

The necessity of providing instruction tailored to the individual needs and aptitudes of each student is the subject of much of the literature on education today. The long-standing goal of providing individualized instruction has not been met on any large scale. It remains one of our chief concerns.

It is only since computer technology has been introduced to the education field that educators see, at last, a means of accomplishing their goal of individualizing instruction.

In the course of this workshop the literature on research in programmed instruction and its principles and applications has been reviewed in some detail. When we received the literature on programmed instruction as a background for the work at Stanford, we found two reasons why so many of the comparative experiments using programmed lesson material found little or no significant gain over regular classroom instruction. First, many of the studies were of very short duration. Overlearning by both experimental and control groups may have confounded the results. The second factor was a result of the way in which the programs themselves were constructed. No matter how many remedial loops a program may have, for the student who makes no errors, the content of that program can be no more than that contained in the linear part of the program. All that can be expected in terms of student achievement is that the student may take less time to "learn" a given body of material than if he were in a regular classroom. There is no reason to expect greater achievement if no more is taught. Remedial loops serve only to bring a person up to an acceptable standard level of performance.

The primary advantages of computer-assisted instruction (CAI) over standard programmed materials lies in the capacity of the computer to keep detailed records on the performance of each student, maintain a large variety of curriculum materials on line, and provide immediate feedback to the student while he is working. In addition, detailed reports of each student's performance can be made available to the teacher on a daily basis.

Levels of Interaction

Dr. Patrick Suppes, Director of the Stanford Project in CAI, has identified three levels of interaction between the student and computer program (Suppes, 1966). Programs for the first two levels have been developed and are currently being used in the Stanford Project.

<u>Level 1: drill-and-practice systems</u>. The drill-and-practice system is designed to <u>supplement</u> the regular classroom instruction by providing an individualized program of review and practice on fundamental skills. This is the simplest of the three levels of interaction. Programs at this level can be prepared and implemented with much less difficulty, and at less cost, in a wide variety of subject areas.

Since this level is the subject of this paper, I will discuss the other two levels briefly then return to examine this first level in detail.

<u>Level 2: tutorial systems</u>. In a tutorial system the aim is to provide complete, or nearly complete, instruction rather than supplementary instruction. The Stanford Project is currently running tutorial programs in logic-algebra and problem solving on teletype terminals. Over the past two years it has developed programs in beginning reading and arithmetic for use on an IBM 1500 system. In the 1500 system first and second grade children are seated at consoles which can project colored film displays on an image projector or graphic displays on a television-like cathode ray tube. Children respond either by touching a light pen to the face of the cathode ray tube or typing characters on the attached keyboard. Instructions are given by an audio system to each child individually through a headset he wears. All

branching and sequencing is handled by the computer automatically while the student responds to problem items as they are presented.

This system necessarily is much more complex than the first level system. The amount of effort required to prepare the program is considerably greater. The hardware cost is also greater.

Level 3: dialogue systems. The dialogue system does not currently exist. It is seen as a system in which free dialogue is possible between the student and the program. Major problems to be solved before this system becomes a reality include speech recognition, the ability of a program to interpret and "understand" a student's question or answer, and the ability to compose answers for questions students may ask.

Work is being done on each of these problems at various centers, and there is reason to hope that within a few years much progress will be made.

A variety of ways to individualize instruction have been suggested. Bloom (1968), Carroll (1963), and Gotkin (1963) have called for programs that would select an instructional sequence for each student based on such factors as verbal, spatial or general reasoning aptitudes, or personality variables such as anxiety or impulsivity variables. Perhaps the work of Piaget on stages of development suggests that programs for students at different age levels ought to be organized and presented differently. White's concept of temporal stacking (White, 1965) would seem to imply that programs written for students younger than six or seven years should be mainly linear, S-R, with very small steps between frames. Programs written for older children might well proceed along current lines since White maintains that between the ages of five to seven years humans shift from animal-like to human-like thinking. Further, there is some evidence that the S-R paradigm for learning applies to some areas for adults throughout life although there is a growing resistance to learning most subjects by this method after the seventh year.

Studies by Hunt (1961), Bloom (1964), and others suggest that programs prepared for use with culturally disadvantaged ought to be more reinforcing, allow more time for responses to be made, and emphasize perceptual rather than symbolic approaches to the development of concepts.

I would like to return now to the consideration of drill-and-practice systems, particularly the Stanford Project's drill-and-practice program in elementary mathematics.

Individualizing Instruction in Basic Skills

How does this program at this level provide individualized instruction? This program of daily lessons in arithmetic fundamentals is intended to review topics two to four weeks after they have been formally introduced in the classroom by the teacher. The lessons are administered to each student individually via computer-based, remote-control terminals as part of his daily instruction.

The content of the year's work at each grade level has been divided into concept blocks or units. Each block contains lessons for seven days' work. The sequence of concept blocks is selected by the classroom teacher to coordinate with the order in which he plans to introduce concepts during the year. The order of blocks may be rearranged whenever necessary to fit particular situations which may arise during the year. Adapting this program to any given text or course of study requires no more than reordering the blocks in the required sequence. To provide for rapid or slow learners, blocks from other grade levels may be inserted in a sequence where needed. A different sequence of concept blocks may be selected for each class. In some cases a "class" may be simply a fast or slow group in a classroom.

The first day's lesson of each block is a pre-test. Based on pre-test performance, per cent correct, the computer automatically selects for each student one of five lessons, each of a different degree of difficulty, for the following day. As soon as a student completes a lesson his performance is automatically computed in terms of per cent correct on the first response. The student is immediately scheduled for a lesson of greater difficulty, the same difficulty, or less difficulty as determined by the per cent correct. The level of difficulty of the lesson assigned each student is a function of his own performance on the previous lesson. A post-test constitutes the seventh and last day of each drill block.

Figure 1 presents a diagram of the structure of a concept block. Each darkened circle represents a lesson. A given student will take only seven

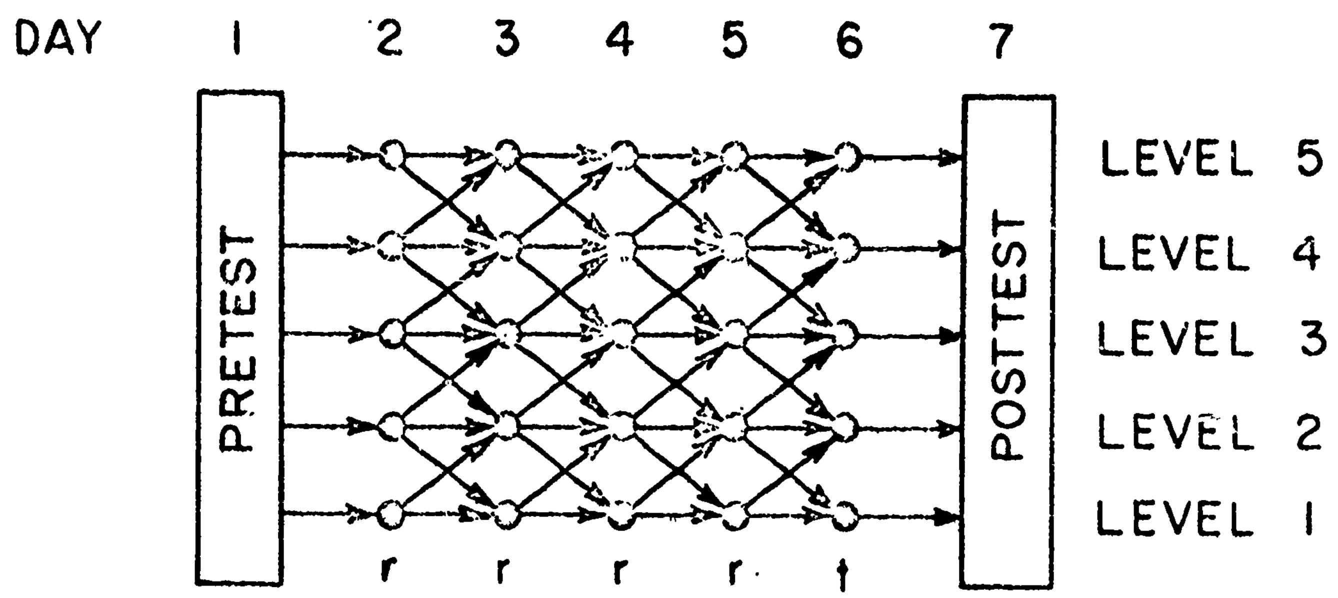


Figure 1. Diagram of branching structure followed in constructing sets of exercises for concept blocks.

of the 27 lessons available in each block. Level I is most remedial in nature and level 5 most difficult. The average student is expected to work at level 3.

Essentially, students are being tested (pre-test) on each concept, such as addition, subtraction, etc., and, according to individual scores, are assigned to one of five math groups each working at a different level of difficulty. Following each lesson, the students are automatically reassigned to appropriate difficulty level groups by the computer. In addition to the regular drill lessons for days two through six, students are given individual review lessons (noted r in Figure I) selected from one of the previously completed blocks in which the student had the lowest post-test score. Each student may be reviewing a different concept, again at an appropriate level of difficulty determined by his post-test score. Following four days of review, the student is given a review test (noted t in Figure I). The review test score that replaces the previous post-test score will be used to determine whether review lessons will be selected from this concept block in the future. The daily lesson in the regular concept block constitutes approximately 70 per cent of each day's work, and the remaining 30 per cent is individual review. This schedule assures that each student will be periodically reviewing his weakest area throughout the year. On the average, from two to ten minutes are required to complete each drill lesson.

Students are expected to take at least one drill each day. Drills include verbal problems as well as most other types of exercises commonly found in popular texts. There are a large number of mixed drills in which exercises are presented a variety of ways. Responses are reinforced immediately.

To initiate a lesson a student is required to type his assigned number and first name. When this is correctly done, the program (the lesson) begins. If an error is made, the student is asked to try again. Figure 2 presents a copy of a sample lesson showing the sign-on procedure and correction routines.

Once a lesson begins, each problem is completely typed out under computer control, including a blank for the response where needed. The typewheel of the teletype is positioned at the blank so that the response will be properly placed. A correct response is reinforced by the appearance of the next

```
PLEASE TYPE YOUR NUMBER AND NAME.
C
PLEASE TRY AGAIN.
143C
PLEASE TRY AGAIN.
143 CON BUTLER


     DRILL NUMBER 8401000


THIS IS A LESSON ON ADDITION.


4 + 4 = _8_

  2
+ 5

TIME IS UP

  2
+ 5
  7

3 + 2 = _5_

  8
+ 6
  5
NO, TRY AGAIN

  8
+ 6
  5
NO, ANSWER IS 4

  8
+ 6
1 4

1 3
+ 3
1 6

5 + 9 = _4_
NO, TRY AGAIN

5 + 9 = ___
TIME IS UP, ANSWER IS 14

5 + 9 = _5_
NO, ANSWER IS 14

  9
+ 8
2 2
NO, TRY AGAIN

  9
+ 8
1 7

9 + 9 = _18

_3_ + 8 = 11

 2 0
+1 2
 3 2
```

```
6 + ___ = 14
TIME IS UP

6 + ___ = 14
TIME IS UP, ANSWER IS 8

6 + _8_ = 14

14 + 5 = _19

4 + 3 + 7 = _14

14 + _7_ = 21

_7_ + 8 = 15

3 6
+ 8
4 4

16 + _6_ = 22

  4 7
+ 8 7
1 3 4

_9_ + 8 = 46
NO, TRY AGAIN

_8_ + 8 = 46
NO, ANSWER IS 38

_38 + 8 = 46

  6 3 8
+ 2 9 3
  9 3 1

     END OF DRILL NUMBER 8401000


20 PROBLEMS WITH 75 PERCENT CORRECT IN 231 SECONDS.


GOOD-BYE, CON. PLEASE TEAR OFF ON THE DOTTED LINE.
................
```

Figure 2. A sample lesson showing sign-on procedure and correction sequences.

exercise. When a first response is incorrect, the words, "no, try again," are typed out and the exercise itself is retyped. A second error on the same exercise is followed by the message, "no, the answer is __," with the correct answer being displayed. The exercise itself is then retyped once more to allow for a correct response. If a third error is made, even though the student has been told the correct answer, the correct answer is given again; but whether the third response is correct or incorrect, the next exercise is presented.

If a response is not given within a predetermined interval of time, usually ten or fifteen seconds, the machine response follows the above pattern except that the words, "time is up," are substituted for the words, "no, try again," at each step described above.

In summary, individualization of instruction is provided by the ability to arrange the sequence of concept blocks, adjust the sequence as needed, select blocks from different grade levels for use by a given class, and provide lessons at five levels of difficulty for each student automatically. Students may work through material as rapidly as desired simply by taking more than one lesson each day. They may catch up following an absence in the same way. By adjusting the difficulty level, poorer students can have successful experiences just as well as students of high ability. In addition, students are reinforced immediately following each response.

Lessons on the lower levels use smaller steps or simpler problems. They also use graphics wherever possible.

Daily Operations in the Classroom

Implementing the drill-and-practice program. Students are told that work on the instructional terminals is part of their daily arithmetic program. Every student in the class is expected to take his turn daily on the machine. Most students like work on the machine and do not have to be reminded when it is their turn.

The attitude of the teacher is a very important factor in determining the attitude students will have toward working on the terminals. A positive and enthusiastic attitude toward the program is reflected by the students.

Students are not told details of the lesson block structure. That is, they are not aware of "levels" as such. We have found it sufficient to tell them that the computer always knows which lessons they are to receive and will be sure to give them the right one. The lesson they receive probably will not be like the lessons given other students, but that is the way it is supposed to be. Everyone is given the proper lesson. The whole class does not have to work the same lesson. Students seem to like this approach. In fact, some students have complained if lessons on successive days have looked too much alike. They thought the computer was not doing its job.

Familiarizing students with the machine. Students have little difficulty learning how to use the machine. The response mode is intentionally kept as simple as possible to facilitate learning and responding. The vocabulary has been carefully controlled at each grade level. It has been our experience that second and third grade children learn to use the machine more quickly than either younger or older children. They learn much faster than adults. Some first grade teachers have prepared dittos of the keyboard of which children practiced finding the letters of their names placing their fingers on the correct location as their names were spelled. Other teachers have borrowed typewriters for a day or so and had children practice in a more realistic setting. In most cases, however, simply explaining the program to the students and having an adult present the first one or two times the student used the machine have proven sufficient.

Establishing a daily routine. The most successful classes have been those in which a regular routine has been established and followed. In some classes a roster of student names is written on the board. As each person completes his lesson, the next follows in turn. This has been used particularly well in cases where there is one machine per classroom. In schools where the machines are clustered in one room, the name tag method has been used successfully. Tags are prepared with the names of seven or eight students written on them. When a student returns to the room after completing his lesson, he gives the tag to the person whose name is next on the list. Whatever method is used, having a well-established routine is very important.

Evaluating Student Performance

Daily report for teachers. A general report of class performance and status will be given daily. This report will indicate areas of strength as well as those which need review. While the format of this report may vary, it will provide information, in terms of class averages, on each of the areas in which any member of the class is currently working. A sample of this report is shown in Figure 3.

End-of-block reports. This is a detailed report of class and individual performance on each lesson. This report will be provided sometime after a class has finished a given concept block. It will be printed by the terminals in the schools as in the daily report mentioned above. A sample of this report is shown in Figures 4 and 5.

Summary

Although the Stanford drill-and-practice program was written for use in grades 1-6, it has been used successfully in junior high school remedial programs. Student response there has been just as enthusiastic as that of elementary school students.

The program has been used successfully with students of all ability levels and social classes. Currently more than 2,000 lessons are provided daily to students in Iowa, Kentucky, Mississippi, and California schools. During the 1966-67 school year, some 79,000 lessons were given. The number of lessons given this year is expected to be four or five times as large. The schools are: Iowa Job Corps Center, Clinton, Iowa; schools centered around Morehead State University, Morehead, Kentucky; McComb City Schools, McComb, Mississippi; and schools in the San Francisco Bay Area around Stanford University. The single computer at Stanford operates all student terminals simultaneously.

Critics of the drill-and-practice level of computer-assisted instruction maintain that the full capacity of the computer is not being utilized. They feel this is an uninteresting, trivial application of computer technology and that more sophisticated things ought to be done.

I would agree that more interesting things can be done at the tutorial level, and we have several programs in operation at the present time. However, the cost is considerably greater and such programs are not ready for daily use in schools. As to the criticism that drill-and-practice programs are uninteresting, I would counter by saying they are also uninteresting for the classroom teacher, hence the source of the original difficulty. It seems to the writer that this is a perfectly natural first step in the development

PRE-TEST

10 MISSED PROBLEM NUMBER 15

```
  4 3 6 5
+ 1 0 7 9
```

7 MISSED PROBLEM NUMBER 17

```
  2 7 0 3
-   4 2 6
```

8 MISSED PROBLEM NUMBER 4

```
  7 6
- 2 3
```

7 MISSED PROBLEM NUMBER 18

```
  7 8 4 3
+ 4 6 7 9
```

POST-TEST

8 MISSED PROBLEM NUMBER 18

```
  7 8 4 3
+ 4 6 7 9
```

6 MISSED PROBLEM NUMBER 15

```
  4 3 6 5
+ 1 0 7 9
```

5 MISSED PROBLEM NUMBER 17

```
  2 7 0 3
-   4 2 6
```

4 MISSED PROBLEM NUMBER 19

```
  8 0 0 2
- 3 4 7 9
```

Figure 5. A Sample End-of-Block Report, Part 2

STUDENT CONCEPT REPORT.

CLASS: 32 GRADE: 4 CONCEPT: 5

NUMBERS ARE IN THE FORM: LEVEL/PERCENT CORRECT

NUMBER AND NAME	PRE		1		2		3		4		5		POST	
777 LAURIE MCCARTHY	C	80	5	100	5	85	5	100	5	92	5	100	A	90
778 COLLEEN MCLEAN	A	80	5	71	5	78	5	71	5	92	5	78	B	85
779 CHARLES SCHULZ	A	85	5	92	5	92	5	85	5	92	5	100	C	95
780 EVELYN BANKS	C	60	4	78	4	85	5	64	5	85	5	71	A	90
784 ANNIE GERARD	C	80	5	57	4	92	5	71	5	71	5	85	A	85
787 ROBERT LASH	A	75	4	71	4	71	4	71	4	100	5	50	B	90
790 SUSAN SMITH	C	90	5	78	5	92	5	100	5	85	5	85	A	95
792 CHRIS BALL	C	65	4	85	5	71	5	57	4	100	5	92	B	90
793 EMILY BREWER	C	80	5	78	5	85	5	85	5	85	5	71	A	90
795 CATHY COGUILLARD	B	70	4	78	4	64	4	78	4	71	4	64	A	90
796 KEVIN HARDIMAN	B	75	4	78	4	85	5	71	5	64	5	64	C	60
806 CHUCK ULRICH	B	80	5	71	5	92	5	71	5	85	5	85	A	90

E N D O F S T U D E N T T R A C E R E P O R T

Figure 4. A Sample End-of-Block Report, Part 1

```
DAILY REPORT          SCHOOL: LAB

DATE: 10 OCT. 1967     TEACHER: DEMONSTRATION -- GRADE 3

CLASS:    3

THESE STUDENTS DID NOT RUN TODAY
?   43 SUSIE SMITH
   110 KAY AAKER
   111 MARY ANNE LANGDON
   116 JANET SMITH   GRADE
   117 GERTRUDE MARGAH
   118 ELEANOR COOPER
   119 IRENE LANCENDORFER
   163 MARGE MEEKS
   164 KAREN JACOBSEN
   187 WARREN FESMIRE   GRADE

STUDENT AT LEAST 20 PERCENT ABOVE THEIR AVERAGE
     3 DEMO N  STRATION                             B302005  75   53

STUDENTS AT LEAST 20 PERCENT BELOW THEIR AVERAGE
?   33 SALLY SMITH                                  L303045  21   65
   162 JANIE FINGER                                 L301015  71   92

STUDENTS WITHIN 20 PERCENT OF THEIR AVERAGE
    13 BILLY JONES                                  L305013  57   68
    23 JOHNNY JONES                                 B302001  75   65

YOUR CLASS IS CURRENTLY WORKING ON:
CONCEPT  302  305  303  304  301
AVG PCT   75   52   67   60   85

THE MOST FREQUENT REVIEW BLOCK WAS 303
```

Figure 3. A Sample Daily Report for Teachers

of an operational instructional system. That is, since no known machine can approach the adaptability and resourcefulness of a human teacher, why not develop a system which can individualize the more routine aspects of instruction and complement the teacher's efforts? The program described above certainly provides more lessons at different levels of difficulty than the normal teacher would ever attempt. It also reports to the teacher daily, in detail, the progress of each student. Using the information provided, the teacher is able to do a better job of individualizing his instruction and taking care of learning difficulties than ever before. At least for the near future it would appear that it is the simpler system, rather than the more complex one, that will be of greater benefit to the teacher.

The Tutorial Approach

The notion that all programmed instruction, whether computer assisted or presented in some other form, must proceed according to well-established principles of learning and programming is no longer held by an serious writer. The once prominent idea that a science of programming could exist independent of subject matter has been dead for some time.

The curriculum material itself determines, to a large degree, the structure of the program. The author's inventiveness together with his knowledge of the subject matter and the user population for which the program is intended are the other ingredients. The author is expected to have a basic knowledge of learning theory and programming techniques and use the technique best suited to the development of each concept. A variety of branching, prompting, and reinforcing techniques may appear in a single program. Each technique should be used where appropriate to maximize learning.

The foregoing remarks simply point up the difficulty of discussing tutorial programs in general. There are many varieties.

At Stanford we are moving away from elaborate instructional hardware systems like the IBM 1500 to simpler configurations which have teletype machines for student terminals which operate either from a local computer or the main computer on the Stanford campus.

I will describe one of our earliest tutorial programs and some of its derivatives.

The logic-algebra program is designed to provide instruction in simple sentential logic and lead into an axiomatic development of modern algebra. This program has been used successfully over the past year with bright fifth and sixth graders. To get a logic lesson, the student simply types "L" before his student number.

There are just four things required of the computer while a student is working through a derivation. First, it examines each instruction given by the student to see if it is syntactically correct and valid. If not, it prints an error message. Second, it performs the operation as instructed by the student and prints the result. It is not concerned whether the step contributes to the solution. The student is free to take any approach to solving the problem he wishes. Third, it compares the result to the desired conclusion. If they match, it prints "correct" and begins the next problem. On certain problems it may give hints, especially when rules have been improperly applied.

A simple example might help here. Suppose the problem was to "Derive Q." The program might present the following premises:

P (1) K — Q

P (2) K

Then the student, knowing that affirming the antecedent (rule AA) would produc the desired conclusion, would simply type 1.2 AA which means "to lines 2 apply the rule AA." The computer would then type out the conclusion Q and print "correct."

Of course, this is a trivial example and was intended to show the mode of student response. Students are free to proceed as they wish taking as many steps as they need to solve the problem.

Derivatives of this program include our new problem-solving lessons in which students are given word problems in arithmetic which have very large numbers in them. These numbers are listed as "givens" rather than premises in the logic program. Students can instruct the computer to add two numbers by simply typing their line numbers and "A" for add. For example, the instruction 1.2 A would add and print the result of adding the number on line two to the number on line one. Again students are free to approach the solution to the problem in a variety of ways.

Every problem-solving strategy used by any student becomes part of the data for later analysis. We are thus able to gather fairly detailed information on the approaches used by each and every student.

The title of this paper was Computer-Assisted Instruction and I have been able to cover only a very small portion of the field. Many efforts are underway and much good work is being done around the country. I have reported on only a small part of the work being done by the Stanford Project. I hope, however, I have been able to give you some glimpse of the potential of computer-assisted instruction as it is today and of its potential.

References

Bloom, B. Stability and change in human characteristics. New York: Wiley, 1964.

Bloom, B. Learning for mastery. Evaluation Comment, 1968, 1, 1-12.

Carroll, J. A model of school learning. Teachers College Record, 1963, 64, 723-733.

Gotkin, L. Cognitive development and the issue of individual differences. Programmed Instruction, 1964, 4, 1-10.

Hunt, J. Intelligence and experience. New York: Ronald Press, 1961.

Suppes, P. Tomorrow's education? Computer-based instruction in the elementary schools. Education Age, 1966, 2, 4-11.

White, S. Evidence for a heirarchial arrange of learning processes. In Lipsitt and Spiker (eds.) Advances in child development and behavior, Volume II. New York: Academic Press, 1965, 187-220.

CHAPTER VI ACADEMIC AND SOCIAL CLASSROOM MEASUREMENT

H. P. Kunzelmann

Introduction

For a long time testing has been used as a tool to evaluate the effectiveness of the teaching process. When a child scores a high mark on a test, some educators believe that they have been provided with irrefutable evidence that he has learned his lesson well. And this evidence is used as proof that the teacher has been successful in her efforts to teach.

In recent years there has been a growing disenchantment with the use of the test to evaluate performance in the classroom. Today, many educators are beginning to question the validity of the test as an assessment of learning, particularly in the primary and secondary school grades. Students in colleges and universities across the country have expressed their dissatisfaction with the examination as an assessment of achievement. Many of our high school and college students have held examinations in such contempt that widespread cheating has become the order of the day.

Testing did not always play the prominent role that it does today. The tutorial system prevailed, with a relationship growing up between master and pupil. Socrates would take his peripatetic walks with his students, and the dialogue between master and students became the educational process. In feudal times the nobleman brought scholars from afar to be members of his household to teach his children Latin and Greek, Music, and perhaps Natural Philosophy. But, with the democratization of higher learning in western Europe, the examination system took a

3

stranglehold. Perhaps one reason was that in many countries of Europe higher education was free. This was the case in England, except for the public schools, Oxford and Cambridge, which, paradoxically, were not. In England, France, and other countries of Europe, the entrance examination was standardized. In each country, those students desiring to enter an institution of higher learning had to take and pass the same examination in order to qualify for admission. In the United States, we too developed standardized tests for admission to colleges and universities. The College Entrance Board examination is required for entrance to many colleges and universities in our own country.

As the population grew, as the standard of living climbed, and as the demand for better educated people with professional skills increased, institutions of higher learning became more selective in their admission policies. High school students desirous of getting into top-grade colleges found that a high score on National Merit Examinations gave them a priority in gaining admission to the university of their choice. The State of New York uses the Regents Examinations, standard tests for high school students. What started fairly innocently has now developed a stranglehold on our educational system. The test as a criterion of performance now permeates the entire educational structure, beginning with first grade and ending only when a person has received his graduate degree.

That the system of testing is ubiquitous in our educational system is a fact. But just because something exists is no proof of its validity. It is high time that we take a good long look and ask ourselves: Does the test have validity as an assessment tool?

One of the principal factors that a teacher wants to know about a student is his progress. She wants to know how much he has learned and how much time it took him to learn it. But most testing procedures conceal this information. Certainly a test can tell a teacher where a student is in relation to his peers in the class, but a test cannot give the really vital information: How effective has the learning process been? The test cannot give any information about what the student has learned when compared with his previous performance nor his rate of learning compared to his previous rate. An alternative tool, and a much more effective one than the test, is a measurement that is concerned with individual role performances.

There has been criticism of educational measurement. A review in a recent publication by Smith and Adams (1966) contains the following remarks about classroom measurement:

> At best, measurements in education are only observation of behavior samples from which we attempt to make emphasis concerning the relative amounts of a quality possessed by different individuals.

Smith and Adams continue their criticism by listing four fallacies in the procedures: One, intelligence and achievement are qualities that cannot be directly observed. Two, man is a poor subject for measurement since he is in a process of continual change. Three, educational concepts, like intelligence and achievement, need to be more precisely defined. Four, the units of measurement have not been precisely defined, consequently the tests that have been developed are not as accurate as dime store rulers.

Most of us would agree with these criticisms about educational measurement. However, most of us would also agree that we need to discover, to develop, and to refine a measurement system that is more precise and more sensitive to change using functional definitions for behavior where there is a meaningful basis for direct observation.

This paper offers a new measurement system for classroom teaching called continuous assessment. Based on the works of free operant investigators, such as B. F. Skinner and O. R. Lindsley, the system is the reflection of work by teachers and university students working in classrooms. Ideally, credit should be given to each person's contribution. At present, this is an inseparable task. In lieu of such credits suffice it to say that teachers, present and past, have made the system functional. Hopefully future teachers will find the system benefits children because of its precision.

Description of a Standard Scale for Measurement of Classroom Behavior

The classroom is a location where an extensive amount of behavior occurs, and yet for a long time it was felt that the measurement of the behaviors emitted in a classroom did not lend themselves easily to objective measurement. However, we have found that this is not so. Most classroom behavior does not occur one time, but many times. A child, for an example, writes the answer 4 to 2 plus 2 many times in his academic life. He writes the letters of the alphabet in printing and in cursive writing many times and more often than not, in the form of words, particularly in his creative writing. The child says many of the phoneme components of words when reading to himself. In the examples above, one factor is common to all: each behavior emitted in the classroom recurs over and over, although under various conditions.

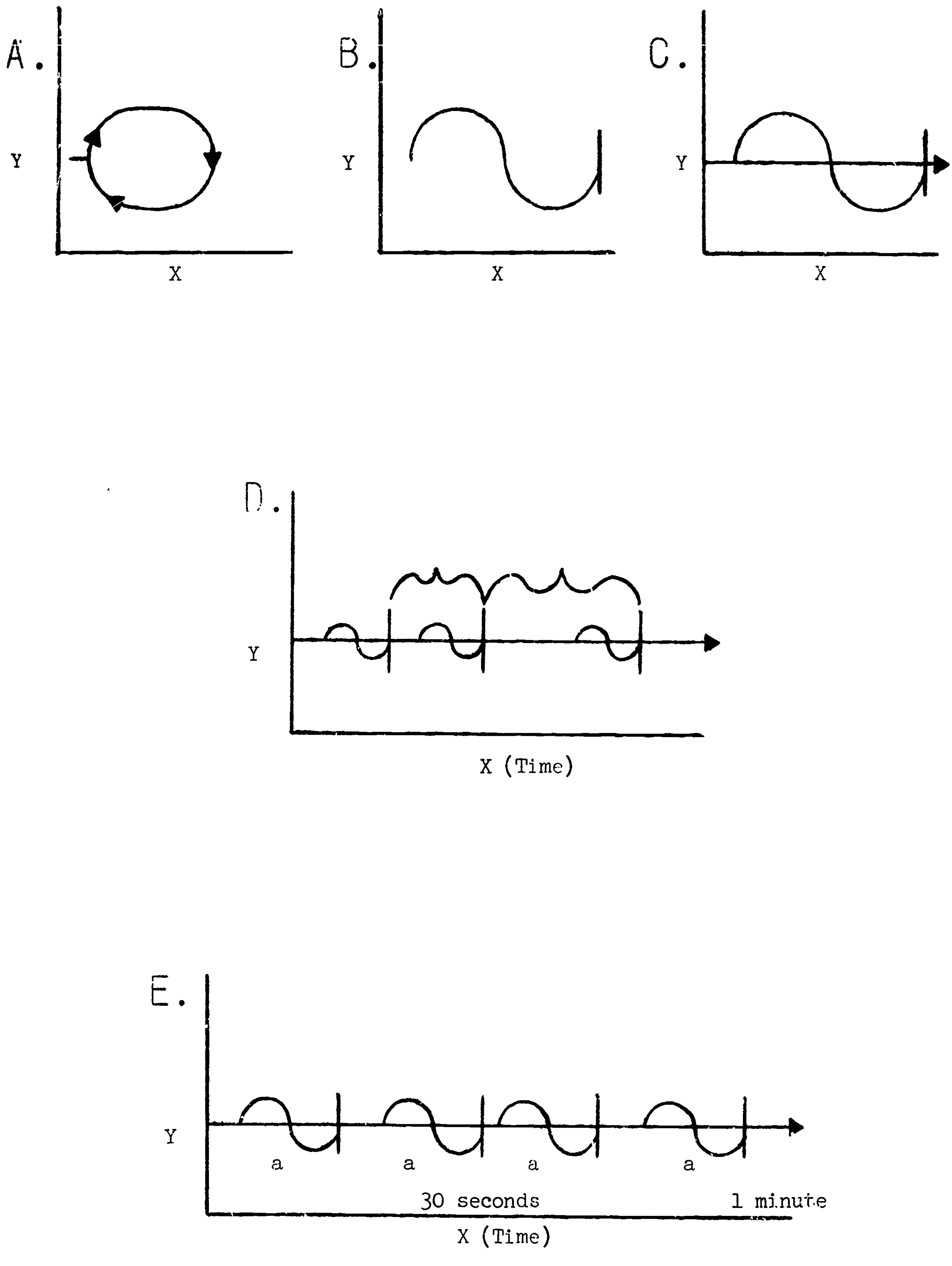
Building a Behavioral Ruler

To devise a valid, standardized measurement system, it is necessary to pinpoint the behavior in a form that can be readily assessed for reliability. Since the measurement form is the actual performance a child emits, such as writing or saying the ABC's, it is possible to obtain a count of the behaviors emitted. It has been suggested by O. R. Lindsley that the term "movement cycle" be used for classroom measurement prior to the functional analysis of classroom behavior. To illustrate, if a student writes the letter A, at its completion he is free to write the letter A again. If he does not write the letter A his emission of the letter A would be zero. However, if he writes the letter A fifteen to twenty times in a given amount of time, information is available about his rate of making the letter A. Thus the movement cycles he emits in the classroom can be clearly defined.

Some of the basic principles about the identification of movement cycles in the classroom evolve from studies of operant behavior. Operant behavior is characterized by certain features, such as, the behavior is usually controllable by the pupil. The behavior always includes some form of movement and is repeatable. Controllable, repeatable behavior that has movement as a component are the requisites of a standard measurement instrument for the classroom. To assess the probability occurrence of any behavior, then it is only necessary to know that a specific behavior is controllable, repeatable, and contains movement.

Figure I illustrates the basis for a description of operant behavior. Section A of Figure I shows a circle with arrows extending from a starting point to the completion of a circle. Starting at the left, moving up and continuing in a complete circle to the starting point is one unit of behavior. Making a circle is a behavior that one can either make or not make, so it is

FIGURE 1.



self-controlled. It is movement, since someone pushed the pencil up, around, down, and back to the starting point. At the end of the cycle, it is replicable because the person can repeat the same behavior.

Section B of Figure I shows the circle broken open and a sine curve shows that the writer, instead of moving back to the starting point, continued around so that a repeatable behavior is shown visually. Depicting the movement cycle in such a fasion visually exposes the movement cycle in time onto the right of the page. The circle could have been drawn quickly, slowly, large, small, heavily, or lightly; however, three factors characterized the behavior: movement occurred, movement ended at a given point, and it could be emitted again.

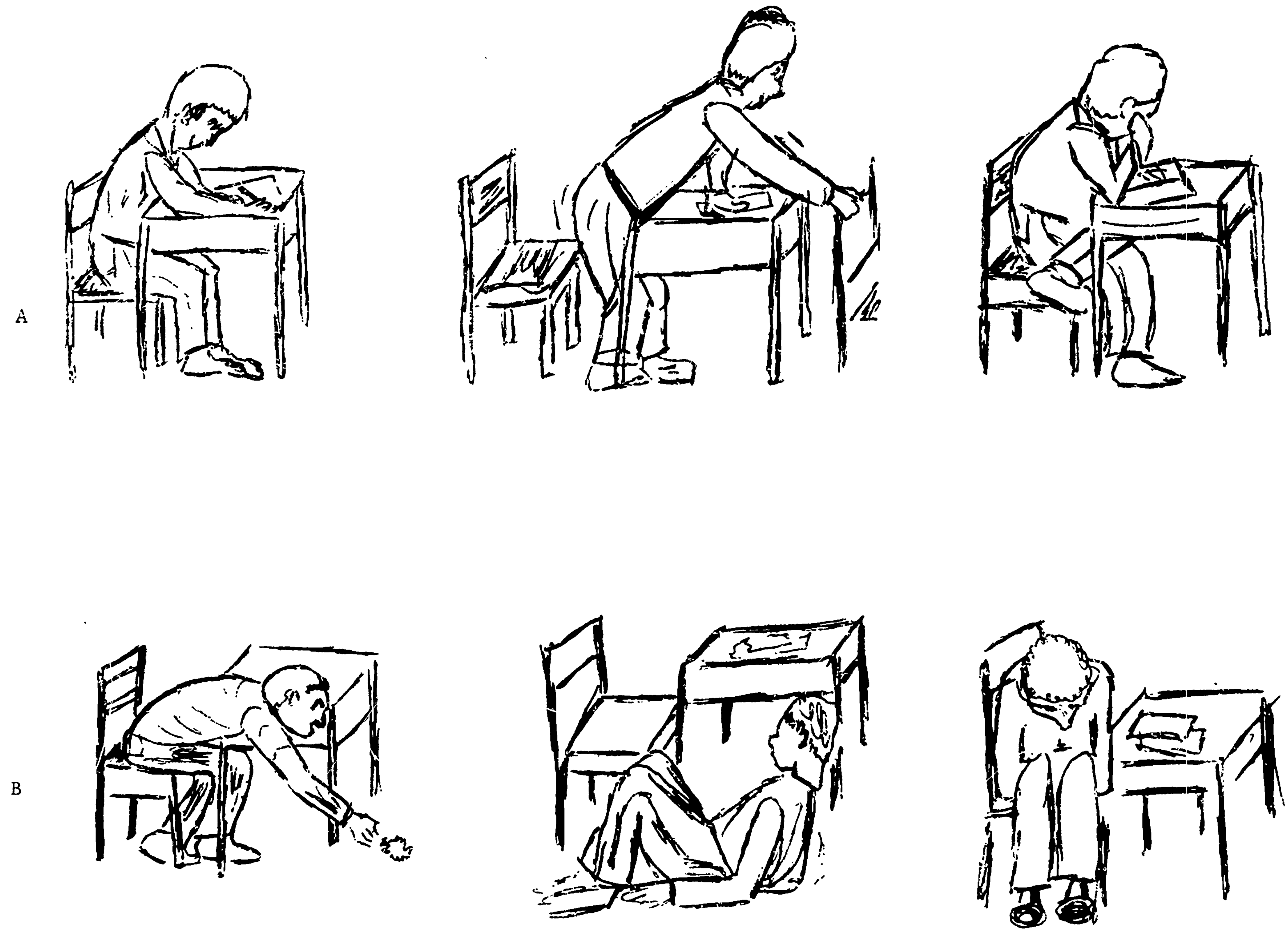
Section C of Figure I also shows that the movement cycle occurred in time along the horizontal or X axis. If one were to make an open circle or sine curve in time, a measurement could be taken between each complete movement cycle. Section D of Figure I shows three sine curves spaced at different positions along the X axis. The distance between the first complete movement cycle and the second is less than the distance between the second and the third. Our standardization would break down at this point if we were unable to use time to assess the occurrence of the movement cycles. By using time and converting it to minutes, we can demonstrate that the movement cycle occurs in a standard unit. First of all, we are using a ratio scale for our measurement system. The behavior may or may not occur. We can use time, which has a zero or measureable starting point and continue for a given number of minutes to observe and count the complete movement cycles.

Section E of Figure I depicts the making of the letter A four times. Subsequently, if the complete movement cycle, making A, occurs in time as shown in Section E, the total movements per minute would be 4. Classroom movement cycles other than writing and saying letters A to Z and writing and saying numerals from 0 to 9, along with various symbols are other movement cycles that can be used.

A typical classroom behavior is that of a child leaving his seat, doing something and sitting down again. Although multiple behaviors occur, from the time that he left his chair to when he sat down again, the measurement is a complete cycle of his out-of-the-chair and back-in-the-chair behavior. To illustrate this more clearly, Figure 2 shows a picture of four ways to leave one's seat in the classroom. There may be many more, but this example should suffice to describe movement cycles occurring with different internal components, where each time a behavior that is controllable, repeatable, and contains movement is depicted. Section A of Figure 2, shows the child as he sits, moves out of his chair, and subsequently sits down; he has touched the chair in front of him. In Section B the child is moving out of his chair, falls to the floor, after reaching too far, and then returns to his chair. Of course, the act of getting out of the chair and falling certainly takes more time than simply leaning forward, touching the chair, and sitting down again. However, in both cases a standard measure can be used since a complete movement cycle occurred. Section C of Figure 2 shows the child in a state of discontent. He leaves his chair, emits a series of behaviors that might be called tantrums, subsequently sits down, and rearranges his desk. In Section D of Figure 2, the child simply leans back in his chair, falls out and again quickly recovers. The common component in all these cases is the movement cycle. The movement occurred, the behavior was complete, and it could be repeated. The teacher may not desire the behavior acted out in these examples, however, it can occur and it can be measured in a standard system from the time the child is in his chair and the movement can occur. Timing of the behavior of the occurrence of the movement cycles starts and stops when the child is no longer in his chair.

The pinpointing of movement cycles in the classroom is essential for the investigation of a standardized system of classroom measurement. It is conceivable that the child doing long division problems will make many more numerals before he is finished with the problems than the child who is doing one column, no-carry, addition problems for each numeral in a measurement cycle. In order to compare a child's performance with his previous performance, or to increase or decrease a given performance in the classroom, standardized measurement must be used. This task is easily accomplished. As an example, Figure 3 shows two graphs of the addition of one column, no-carry arithmetic

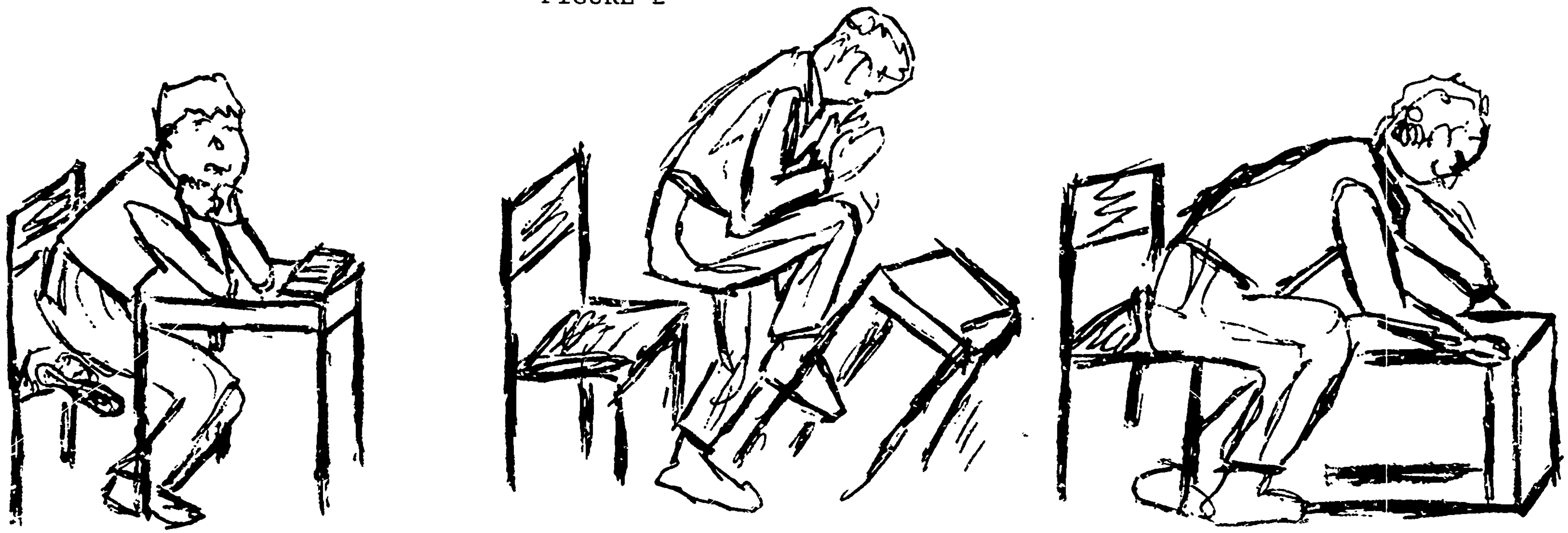
FIGURE 2



Illustrations by Mrs. Ann Mingo, teacher, Experimental Education Unit

FIGURE 2

C



D

illustrations by Mrs. Ann Mingo, teacher, Experimental Education Unit problems. In Section A of Figure 3, it is evident that the fourteen-year-old child emits under three different conditions--working on one-place, two and three-place, no-carry, addition problems--and that the emission of the movement cycles in composite form reduces his performance rate. This also holds true for the performance of the thirty-one-year-old male subject. Although the performance rates are higher, the drop and the direction seem to be a result of the way the performance was counted. The teacher, in measuring the performance, must count all the numerals involved if an exact measurement is desired. Although this may not be necessary, a teacher having information about movement cycle grouping and knowledge of the component parts can then isolate possible causes other than teaching functions for the data or the measurement to change.

Using a Behavioral Ruler

For the teacher who has been using movement cycles and who has accurately pinpointed children's performances, measurement errors may be of little concern. However, most of us make errors in our classroom measurement related to movement components, to cycle completion, or to a time basis. The readily identifiable errors are those of counting and timing. The importance of avoiding measurement error is reflected in at least two ways; first, a counting error may direct attention only to part of a movement cycle leading either to the reinforcement of the wrong component of the movement or the reinforcement of just a part of a cycle.

Second, timing errors are apt to furnish us with inaccurate information, such as counts that fail to encompass complete cycles, and therefore curtail one from using various schedules for reinforcing.

Figures 4, 5, and 6 illustrate the various errors to be avoided. Figure 4 shows the counting error potential when movement cycles are incomplete. The frequency of occurrence of a behavior is basic to counting. For example, the counts in Figure 4 are not standardized to represent the same units. There is a lack of completeness to each cycle. A specific standardized count may be thought of as a behavior such as a button push. Pushing a button down cannot be repeated until the button comes back up. Generally, a standard amount of force is needed to push the button down, the return

Figure 3A

SEPARATE INSTANCES OF OPERANT VS RESPONSE GROUPING

14 YEAR OLD SUBJECT

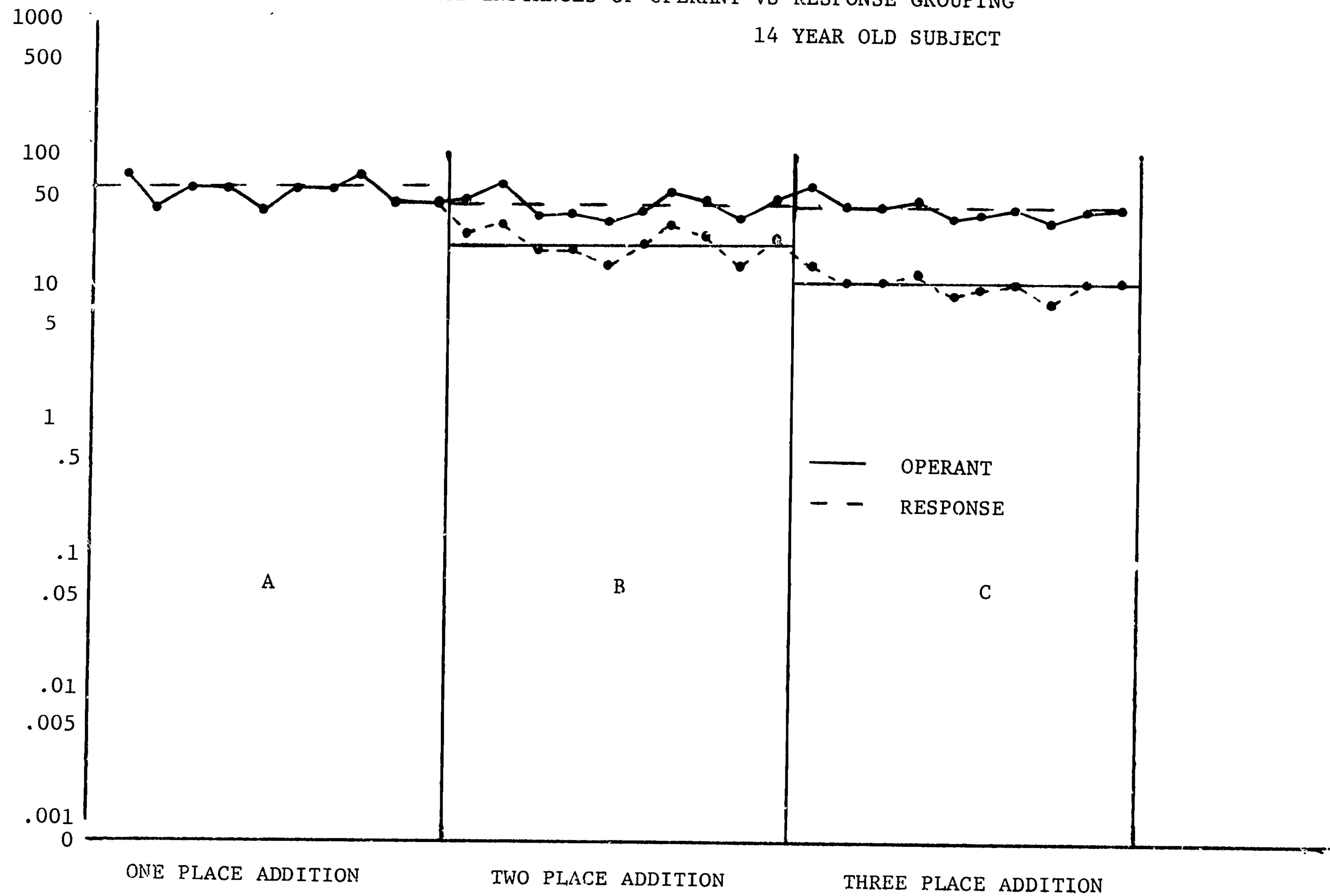


Figure 3B

SEPARATE INSTANCES OF OPERANT VS RESPONSE GROUPING

31 YEAR OLD SUBJECT

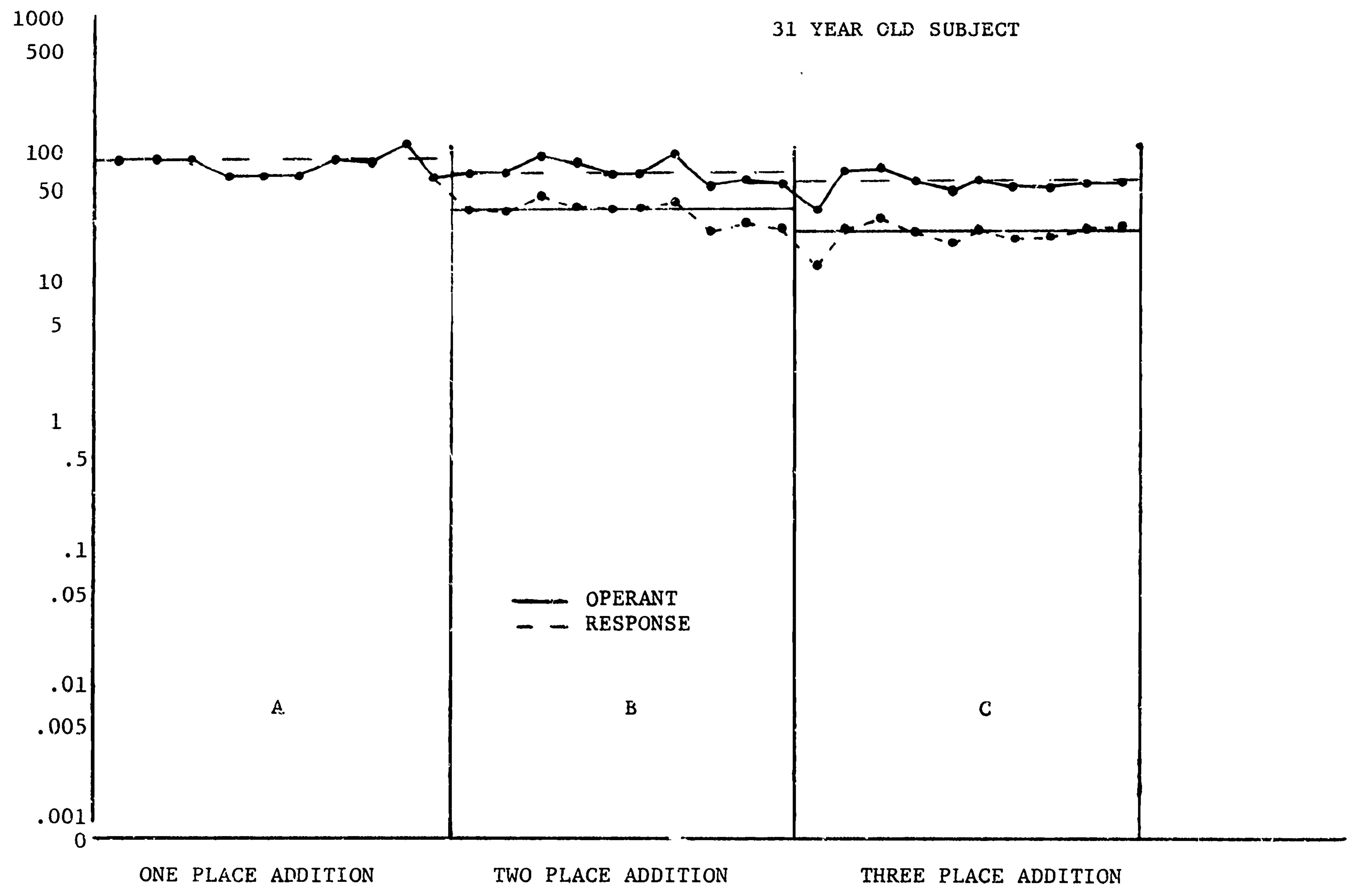


FIGURE 4.

Counting Error Risk

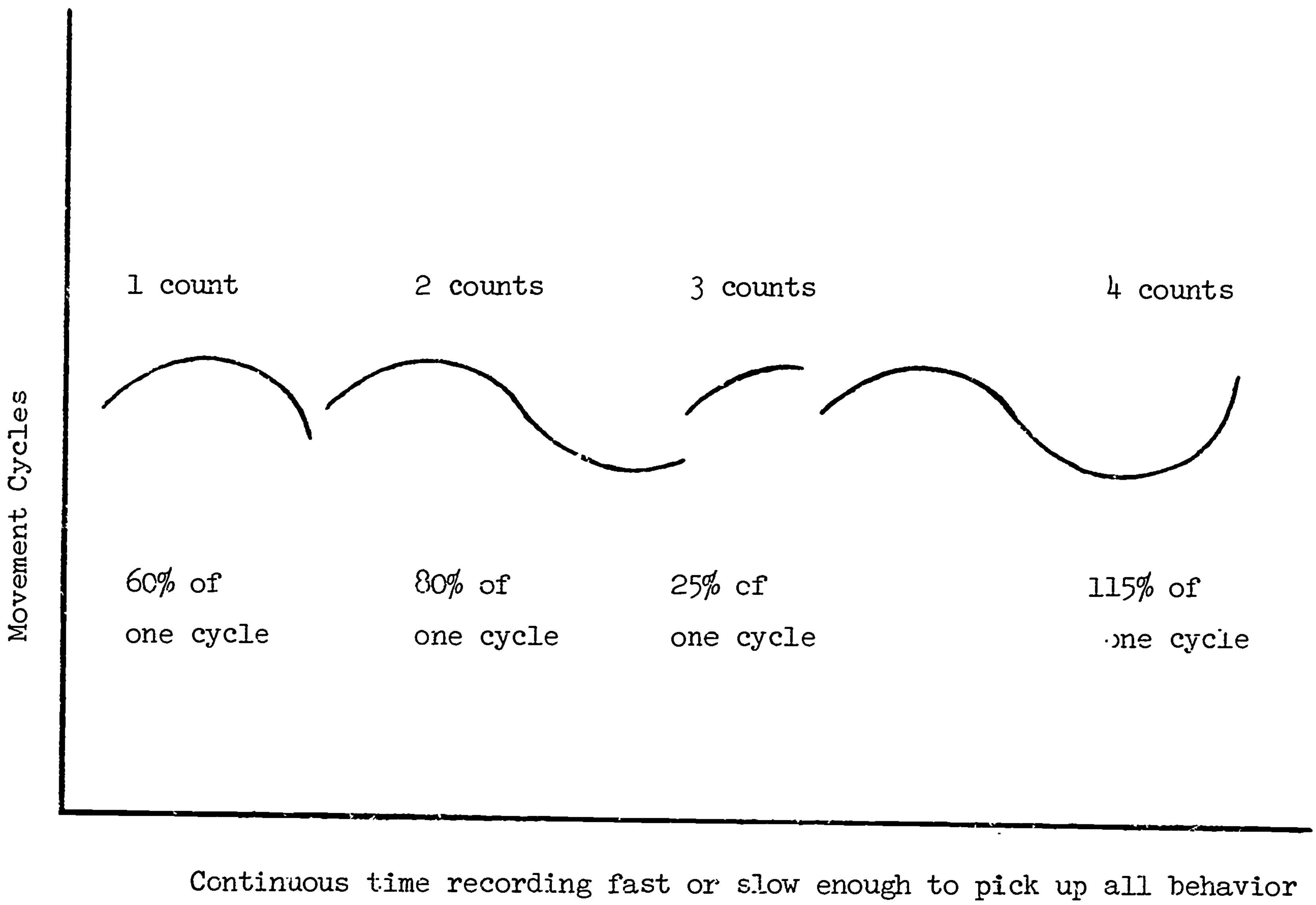


FIGURE 5.

Timing Error Risk

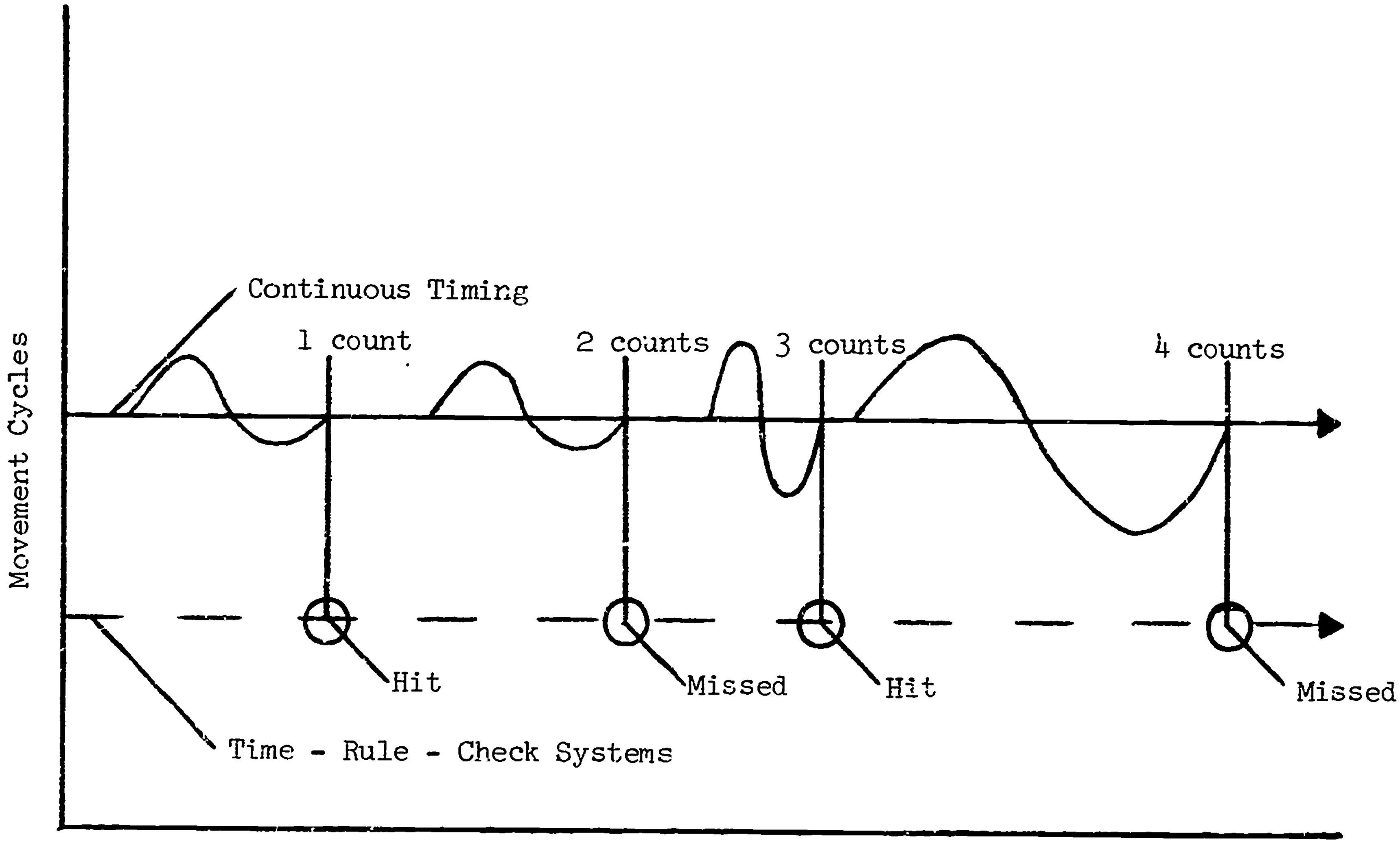
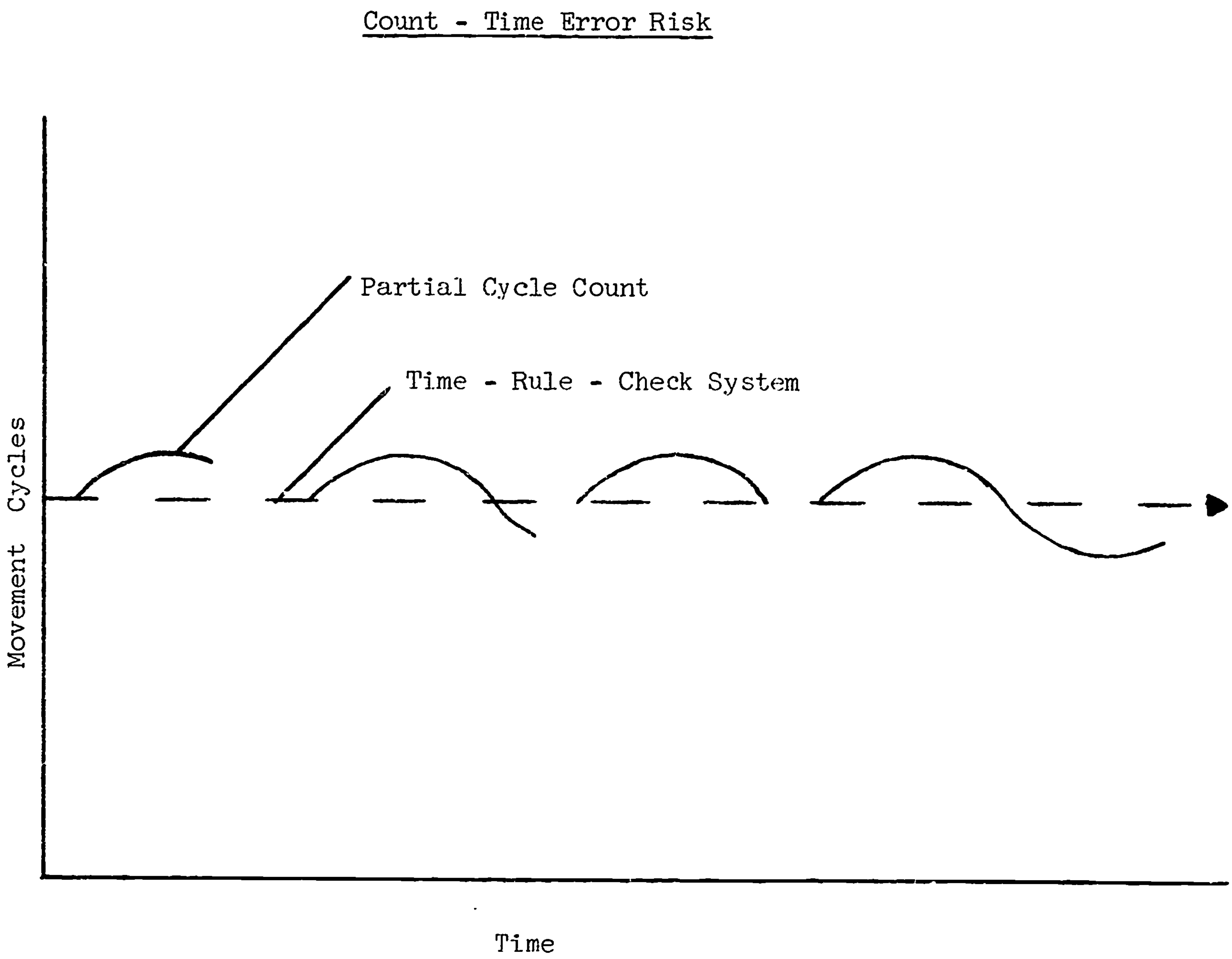


FIGURE 6.


Count - Time Error Risk
Partial Cycle Count
Time - Rule - Check System
Movement Cycles
Time

movement is simply a springload. The counts are dictated by the responding device and the movement cycle. In the same way, we try to find equal units in writing and in speaking. The desired letter is made so as to ensure that others may read it. The child can match or correct himself and his rate will increase with use under multiple classroom conditions and curriculum exercises.

Figure 5 shows the basic timing error risk when movement cycles are complete, repeatable counts. The sample timing allows for hits and misses on the count. Time rule checking, as a recording technique, has limited functions for several reasons; first, one can note from Figure 5 that even when the movement cycles are complete, reliable counts, some count points may be missed. Second, only under conditions of known complete movement cycles of fixed duration can the sample ystem be measured reliably.

An excellent example of a timing error can be found where Kunzelmann (1967) described a five-second sampling technique showing simple pacing but clearly limiting the amount of data and distorting reliability measures by crossing movement cycles.

Figure 6 depicts the combination of observation and recording errors with partial counts and time-ruled check list locked. Such errors which are not attributable to the recorder for inter-rater reliability could be high, but they are meaningless between cycles.

Included in a discussion of potential measurement error risk is the reporting of the time unit carrying the count. Time units usable for frequency counts are seconds, minutes, hours, days, weeks, months, and years, to name those most often used. Classroom information usually is given in terms of gross time units, such as, Johnny's quarter grade was a B in English or Johnny has done well this past week in math. Although the pinpointing of the movement cycle is obviously lacking in these statements about Johnny, the issue of time distorts each statement. Comparisons of the two measures would not be possible unless either one week's information were projected into nine or ten units, according to the length of the school year quarter, or the nine units were averaged into a weekly statement. Rather than such gross reporting, it is preferable to use one unit where most information has the same scale. Minutes serve this purpose most readily. Minutes as a base point for human movement cycles encompass a wide range of classroom activities. Schools generally have functional hours; for example, 8:30 a.m. to 3:30 p.m. each day. Programs or daily lesson plans, however, are expressed in terms of minutes such as: 50 minutes of math, 40 minutes of social studies, 10 minutes of spelling, to name a few. Rarely are seconds or months used in the lesson plan. At times, weeks are used for specific purposes such as, "we will cover a unit of social studies for the next week," but not used as time units in the presentation of subject matter.

To ensure that data has comparative possibilities, the time unit of minutes seems to be most appropriate. A child may read orally at the rate of 100 words per minute; write numbers at 50 numbers per minute; say words at 110 words per minute; talk out in class at the rate of .01 per minute, or once every 100 minutes; make errors when writing answers to math at the rate of .5 per minute, or 5 times every 10 minutes; and hit a peer at .002 per minute, or 2 times every 1,000 minutes, while being observed for 427 minutes or 7 school hours. A case for using minutes as a base for expression of classroom counts evolves clearly by description.

Errors in time units may not always present a measurement problem. Potential error in classroom measurement may also occur when different unit analysis possibilities are reduced. A teacher who has used continuous assessment measures in a classroom may want to determine the multiple effects of a given instructional procedure for one child. Rate statements of many performances, from pinpointed movement cycles, are of little value if the time units are unequal.

Let us consider what problems a teacher may find. First, he may plot the data per day and find that rate of performance per day shows that a child ranges between .5 correct numbers per minute and 18 correct letters per minute. At the same time, if the teacher counts the child's rate of verbal contributions to class discussions per day, comparisons between numbers, letters, and verbal contributions or answers are impossible since gross minute statements can be made. A specific example would be that possibly Johnny participated five times by answering questions in a classroom in 420 minutes. The statement is inaccurate. The programmed time or the number of minutes for possible presentation in the classroom answers could not have been given all day; hence, the time available for the performance would be misconstrued. The measurement fails to account for time for recess, drill exercise, and lunches. The teacher, by using two different time units, minutes and hours of one day, has distorted the picture of the child's performance and cannot make comparative analysis statements about him. At this point, consider briefly what one would do with performance rates on many children if, at the same time, one attempted to analyze multiple bits of information of classroom data. It is an impossible task to accurately depict the effects of teacher performance in general when such measures distort the actual time factors in performance.

To avoid the error risk in counting and timing and the potential error in analysis of unequal time units in observation and recording, to standardize classroom measurement, and to have a basic behavioral ruler, a complete movement cycle must be used as a standard unit for observation and counting. A behavioral ruler, or standardized unit, is not new to experimental psychology or education in many settings. Such standardization gives the rate of performance in the classroom without concern for compounded measurement errors.

How the Behavioral Ruler Refines Educational Measurement

Earlier we stated that there were at least four general criticisms of educational measurements. The first was that much of the information that is used in the classroom cannot be measured directly. To counter this argument, the suggested movement cycle using a minute-count system for rate of performance in a classroom dispels the belief that there is any child behavior that cannot be placed into a movement cycle format and timed in the classroom by the teacher, the pupil, peers, or other outsiders when necessary. There is a direct measurement potential when through appropriate semantics we clearly identify a specific behavior by ascertaining whether there is movement to the behavior, if it has a cycle, and what the cycle is. We know that we can make a count for frequency information and we also know that we have a measure that consistently gives us the probability of behavior occurrence by using minutes as the standard time unit.

The second criticism of educational measurement is that basically many of our measurements cannot be accurate because man as a subject of study is always changing. Hopefully, the reader is aware that even though man is in the process of change, his behavior can be measured. If skills develop by practice, one must then agree that the effect of practice is measurable in terms of movement cycles that are repeatable and whose frequency can be easily measured when the time unit is small enough to observe the repetition. This rate statement can account for man's change and especially a child's change from stumbling through the Dick and Jane series at about 10 words a minute to reading Dick and Jane No. 11 at 100 words a minute without error. Man's change can only be assessed with a system that allows for the change and accurately assesses it. This is the reason that rate is used. Rate is a simple statement of the number of movement cycles over time where the time is expressed in minutes. The examples given describe the desired scale for classroom measurement. Assessment should not be limited to these, however, but should be attempted with any other performance in the classroom.

The third criticism, pointed out at the beginning of this section, was that precise definitions of educational objectives are needed as a basis for communication. The information concerned with movement cycles allows educational objectives to be precisely defined; precise not only because they can be measured over and over by the same person, but because they can readily be transferred to the next person managing the behavior of an individual. In a classroom, as a child moves from the first, second, third, and fourth grade, very little information about his rate of performance is transferred from one teacher to the next. Miss Smith may know that the child has done well under Miss Jones. She is not sure, however, of how the child will perform when he gets to her class. There is a precise definition available for every behavior in a classroom when movement cycle criteria are used.

The fourth criticism is that we need better tests. Contradiction is at hand. In the introduction to this paper we stated why the test as an assessment of achievement lacked validity. The system for standard measurement in the classroom presented here regards validity as simply a crutch that has been used for the past two hundred years to ensure and rationalize the use of mental tests where movements could not be specified.

The physical sciences have not been concerned with the issue of validity. At any pont where a more reliable measure was found in the physical sciences, validity was immediately disregarded as an issue. If one is measuring accurately from moment to moment, as in continuous assessment, the validity issue is nonexistent. It is suggested here that instead of using tests for the assessment of child performance that we substitute continuous measurement of educational objectives in terms of movement cycles under various conditions set up by the curriculum. Continuous assessment is a tool whose reliability is unique. It can be utilized for any individual at anytime. Continuous assessment has the further advantage of individualization, since the program is adapted to the individual child rather than adapting the child to the program.

Utilizing the Standard Measurement Scale for Quality Classroom Management

In the previous section a description was given on a standard measurement scale, Namely, movement cycles can be directly counted on performances of children in the classroom and that the count of the movement cycles can be placed over time, measured in minutes, to produce a rate measure. The performance rate of any given child is as sensitive as one could desire to determine effects on a child's performance as a function of teacher, peer, or self-produced changes in the environment. The use of the standard scale, or preferably, a behavioral ruler, should be considered only in terms of first, what it will allow us to do differently in the classroom to help children learn; second, what information different from previous information will evolve in a manner that will allow for immediate and individualized analysis; and third, what information can eventually have a predictive function because of the standardization of the described measures.

It was found that the initial step of recording in the classroom had to include the basic datum for the study of behavior. This is rate of movement cycles. It was determined further that it is the child's performance that must be changed. As an example, the child hopefully becomes a reader, where he starts as a child not having the skills and movement cycles to read. He hopefully learns to write letters, where he starts as a scribbler. The child learns to communicate to other children, where he starts as one sitting silently in a corner. The actual behavior of reading, such as saying words, pointing to words, pointing to letters, making letters, is the basic content of how a standard measure is used.

The capturing of each act of a child's performance has been established as a basic measurement unit and becomes a frozen record that the teacher can look at from day to day to make analyses of the gains made by the child and to determine changes and make decisions about the environment that may help the child to learn better. Gains in performance rate would be identifiable as the child starts to read one word; the child now reads two words; he now puts the words together in a sentence; he reads them from a flashcard; he reads them from a book. At the point where the teacher can specify what he asks the child to do, and the words that the child needs to say, to write, or to point to for an appropriate response, quality performance in the classroom is established.

The use of the standard measure to assess pupil performance provides the classroom teacher with an unprecedented tool for determining the effectiveness of his professional skill. Rate of performance, when analyzed from movement cycles or repeatable behaviors, removes all shadows of environmental influences, including inter-pupil comparisons. The teacher, using the tools mentioned, is free to arrange learning environments, to explore them, and to change them to ensure accelerated or decelerated pupil performance. Further, the teacher is now freed of all of the antiquated bonds reflected in such terms as ability, aptitude, intelligence, to name a few. The limitations of the pupil now become the limitations of the environmental arranger, and in the classroom, this is the teacher.

A switching behavior is described which places the learning responsibilities on the environment rather than on the child and will probably have the greatest impact on the children most affected by past error, those exhibiting retarded behavior. The standard behavior ruler will produce new horizons for the investigators. The teacher, in her "find-and-make" environmental world called the classroom will have considerably more opportunity to produce valuable information for other teachers as well as to help the children in her charge. The person, parent, professional, or peer will find that his relationship to the child is quantifiable rather than a semantic explantion of unknown dimensions. The use of rate measures in the classroom places the responsibility on the environmental manager, and subsequently demands that a clear precise communication exists between the pupil being taught by artifical and actual environments and the related manager.

If the classroom consists of a place designed for pupils to perform or recite, then the classroom's environment must consist of those elements necessary to stimulate and consequate the performances of the pupils. The measurement system thus far described not only standardizes the unit measure for the pupil performance, but enables the teacher to know precisely when the manipulation of any part of the classroom environment has been effective in changing the child's performance. Within the classroom environment the teacher is typically the manager, or the person as close to the behavior as anyone else, except for the child himself. The teacher's task can be categorized into two continual goals or objectives.

The Teacher's View of What Must Be Measured

First, the teacher is confronted with a find situation. A find situation says exactly what the word implies. The teacher has a new pupil. A teacher has thirty new pupils. A teacher has a new set of materials for math. A teacher has a new basal series or readers. In each situation the teacher is confronted with finding what effects the environment has on the pupils. The second goal or objective of the teacher is what might be called a make environment. The make environment is one in which the teacher has all the components, builds them into a program with a set of antecedent events, and an arrangement for presenting subsequent events. Under the conditions of both find-and-make environments, acceleration and deceleration targets are readily ascertained. The teacher may desire a zero rate of talk-outs in her class. She may desire a rate of eighty words per minute for oral reading at the fifth grade. She may desire a rate of .1 out-and-in-seat movement cycles. If all of these pupil activities were planned within a make environment, the teacher would have immediate access to acceleration and deceleration information; however, this is not generally the case. The teacher has a new pupil, and the pupil is out-and-in his chair at the rate of .9 a minute, or nine times in every ten minutes. A teacher may find that the target is one of deceleration or he may want the child to perform differently. To be confronted in a continual manner with a find-and-make environmental situation, the teacher can utilize her standardized behavior ruler called movement cycles, expressed as rate of performance, in a most efficient manner.

When using the rate of performance in the classroom, one has only to determine how to count and time the performances of the children. To utilize our ruler, various tools are available that have been tested. The teacher or child should be able to use most of the tools available. If the lesson is teaching basic measurement in the classroom by using a foot ruler, the text, usually a workbook, is available, along with a vast amount of teacher information which will help pinpoint for the child what the ruler looks like, help the child to discriminate the inch marks, the half-inch marks, the quarter-inch marks, the eighth-inch marks, and, in some cases, the sixteenth-inch marks. The teacher will let the child practice by measuring various distances on paper and various objects in the classroom until the child utilizes the ruler accurately. Upon arriving at a standard behavioral measure in a classroom, the same information for measurement must be available for the child and his peers, as well as the teacher. Subsequently, the use of movement cycles as the basic unit of measurement in the classroom enables the teacher to instruct the child and his peers how to count and time behavior.

Continuous assessment in the classroom will probably only be feasible under conditions where the child has been taught to measure many of his own performances. Basic information, such as the number of words per minute, the number of phonemes per minute, and the number of letters written per minute can all be counted in and timed by the child and his peers. Eventually, the teacher will not have to spend many hours of counting and plotting for a child if the child has been taught at a very early age to measure his own movement cycles. Some of the tools readily available and tested for this purpose are wrist counters used by golfers to count the strokes they make on the golf course (all complete movement cycles, by the way), knitting tallies that fit on knitting needles but also fit very nicely on pencils used by children and by teachers, hand tally-counters used by coaches in the gym, pieces of ruled paper to enable the child to make counts and subsequently tally the information, specially designed recording sheets that allow specific information other than the movement cycle count. All these tools are immediately available for utilizing the standard behavioral ruler in the classroom.

Although these tools for counting in the classroom may seem trivial, they are extremely important in that where measurements need to be taken, they should be communicable.

Timing techniques, readily available for classroom use are headed by recording from the classroom clock, the start time and stop time for a given performance. In many cases, this is quite simple. As an example, to record the rate of doing cursive writing, it is quite easy to write the letters of the alphabet continuously for one minute. The start and stop time is simple, from 0 to 60 seconds lapsed time for the time measure of one minute. The count of each complete movement cycle within the minute would be a count of the number of the letters of the alphabet produced. This would mean, for this author, a rate of 104 cursive letters per minute. For each individual the differences will vary according to his reinforcement history for writing. Where time factors vary in terms of minutes, as in most classroom activities, the teacher may want to record the time on the pupil's worksheet. If a drill exercise is being recorded, the teacher may want to produce for the children a system whereby the child records on the top of his worksheet in the lefthand corner the time just prior to starting his work and in the lower righthand corner the stop time of his performance.

Although most timing techniques are simply innovations by teachers, it is not remote to hope that eventually publishers will give consideration to timing performances on the actual worksheets before and after questions at the end of sections in textbooks and that they will include specific suggestions in teacher manuals related to pupil texts. The programming of time units has not been investigated adequately at present. For example, making letters for 10 minutes at 100/minute may cause a tremendous rate drop for other than professional writers who have bread and butter for self and family as a variable consequence. Such issues may, however, be investigated by using movement cycles as the standard measure in the classroom.

Conclusions

The points made throughout this chapter all hinged on the description necessary to outline a standard measurement for recording pupil performance continuously in the classroom. Certainly, a new approach for evaluating teaching is needed and movement cycles as a behavioral ruler seem to hold the answer. What has been obviously lacking in the discussion is empirical data to support all descriptive statements. This is an area of concern and clearly a direction for more data. However, it should be noted that most measures have degrees of exactitude. For example, consider the physician viewing a victim of an accident. Visually, palpitation may be noted in the chest area--one brief but vital measure. Secondly the physician may feel and time the pulsation at various locations on the body. Third, if available, a stethoscope may be used for more detailed measures, and fourth, if available an EKG may be used for extensive measures.

The system of measures cited above range from the very crude to very sophisticated. Each hopefully has a function for saving life. Such gross to minute measures are not restricted to saving life. The lathe operator may use his tape measure for cutting a large block of steel, but when the steel is being milled the same operator uses micrometers and calipers for his measures.

The teacher has the first crude measure of correct and error in most subject areas. Movement cycles measured in minutes is a second step in the classroom measurement system. With more descriptive information, such as how chronological age limits the rate of writing and saying A-Z, 0-9, and a few symbols, we may find need for EKG, micrometers, and calipers where now we use a ruler.

Viewing the Learning Process

Webster's describes a teacher as one who teaches or instructs. Certainly none of us would disagree with that definition. Yet it fails to embrace another important function of the teacher in his professional role.

The teacher is responsible for a learning process. It is not enough to teach a class of children; the teacher must see to it that the children learn. This is not an automatic occurrence. A classroom comprises a complex population. A typical classroom may have about 30 children, and, in a normal set-up they will probably be, for the most part, in the same age group. But that is where the homogeneity of the class ends. For these 30 children may present at least that many academic and social behavior patterns. Johnny may read at 50 words per minute while Susan reads at a much slower rate. David does 2-place addition problems seemingly with no effort while Linda is incapable of completing the same assignment, and even when she manages to write the answers to half the problems in a given amount of time, many of the answers are incorrect. There is the child who habitually spends a good deal of his time out of his seat, or he doodles or looks around instead of writing the assignment.

The teacher is responsible for a learning situation for each child in the class. It is not a matter of Susan learning to read as rapidly as Johnny. Susan's reading progress must be determined by her ability to read at a faster rate what she now accomplishes in a slow and faltering fashion, and when she has achieved that objective, she must be reinforced to go on to more complex reading materials.

If the teacher is to help the child achieve a goal that he cannot attain when he first enters the class, the teacher must have specific information of the behavioral performance--verbal, gestural, social, and academic--of the child. He must also know the probability of the occurrence of each child's behavior in the class. For every behavior to which the child responds, the teacher, using a behavioral ruler, can obtain a measurement. This measurement will not only provide her with a rate of occurrence of a behavior. If the measurement is taken consistently, the teacher will also have a probability statement of the occurrence of that behavior.

A simple behavioral event would be the writing of numbers or letters (movement cycles) or saying numbers or letters aloud (also movement cycles). It is possible to ascertain the probability of these movement cycles when conditions are held constant by the use of a behavioral ruler. For example, Billy, a student in the class, working from 9:30 to 10:30 completed 15 math problems in the fourth grade math book X on page 10. He made 30 numerals over a period of 10 minutes, a rate of 3 numerals per minute. This "3 numerals per minute" is a statement of his work for those 10 minutes. The student of human behavior, however, needs information about possible events in relation to their probable occurrence. Movement cycles that occur within a specific period of time can be expressed as a rate of performance. But what about tomorrow and the next day and day after that? In order to obtain a statement of the probability of performance, continual measurements each day are necessary.

Before a discussion concerning the application of the behavioral ruler, it would be worth while perhaps to consider what requirements it must possess. First there is the area of possible behaviors. A human behavior may occur once in 1000 minutes to 1000 times in one minute. For example, let us use the environmental condition of the infant in a crib with the infant in an awake state for approximately 16 1/2 hours. If a movement cycle of one kick is noted, the rate of kicking can be expressed as .001 per minute. Any number of other examples can be used. The behavior observed could be one throw (of a rock, a chair, a piece of chalk), one yell, or one writing response in 1000 minutes, making the rate of performance .001.

In the opposite direction from one movement cycle in 1000 minutes is the behavior of 1000 movement cycles in one minute. A tool that can be used for observing human behavior has to be able to record the whole panorama of behaviors. Just as the lens used in the telescope enables the astronomer to view events taking place at remote distances from the earth, and in the microscope enables the biologist to examine the infinitesimally small objects on a laboratory slide, a tool for viewing behavior demands the same latitude.

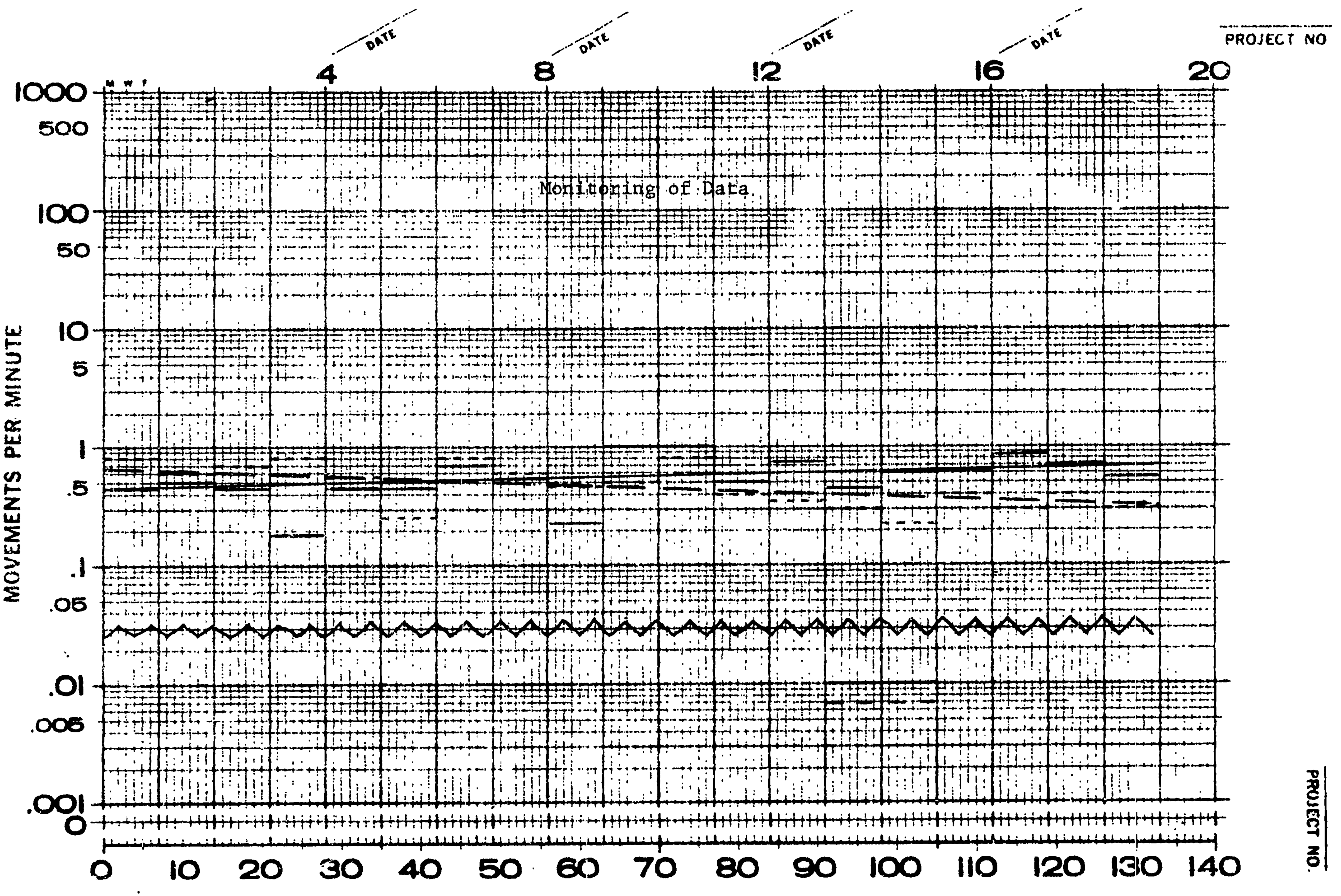
There are movement cycles that occur 1000 times in 1 minute. Saying sounds is such a movement cycle. Phoneme production for most adults may even go beyond 1000 responses per minute. However, at the Experimental Education Unit of the University of Washington, teachers have found that in classrooms the measures generally run from 600 to 800 per minute. Movement cycles occurring at 100 per minute to 300 per minute take place constantly in our daily lives. Steps in walking and running, words in oral reading and letters in writing all fall into the high rate category of our repertoires. The teacher may desire to check movement cycles under various conditions. Many of them will range between 1 movement cycle in 1000 minutes and 1000 movement cycles in 1 minute. It is obvious that any tool employed for recording human behavior must encompass all the movement cycles that may occur.

Another requirement of a viewing tool is concerned with the probability of a behavior. When a movement cycle has been pinpointed, and the frequency of its occurrence has been timed in units of minutes, a measurement of the rate of movement cycles can be obtained. In order to predict the probability of the rate of performance, information must be forthcoming about the rate tomorrow, the next day, the next week. To be adequate for recording performance rates, the tool for measuring performance rates must be able to show performance rates that are continual and that can spread over long periods of time. The viewing tool must have time units of an ordered sequence if the student of human behavior is to be afforded the means of making probability statements.

To be glib, although many tools have been tried, only one has been chosen, which is just a way of stating that a tool for viewing measures of human behavior has been developed. Six-cycle semi-logarithmic chart paper, Figure 7, developed by Dr. O. R. Lindsley and his students, fulfills the tool requirements discussed.

As can be noted in Figure 7, the scale of charting movement cycles extends from .001 to 1000. This meets the requirement of possible human behavioral emission and the time units across the lower edge of the chart

Figure 7



TRAINER	Hulten ADVISER	Mingo MANAGER	Male PROTEGE	19 AGE	LABEL	Write Letters MOVEMENT

hold real time for the distnace of the page (standard 8 1/2 x 11) of 140 days or 20 weeks.

Thus, the telescopic and microscopic aspects of human behavior can be visually depicted. The semi-log paper has, besides meeting the obvious requirements, other advantages. The most obvious advantages have been listed by The American Society of Mechanical Engineers (1960). The six-cycle semi-logarithmic chart:

a. Presents a picture that cannot be shown on an arithmetic-scale chart.
b. Converts absolute data into a relative comparison, without computing.
c. Shows the relative change from any point to any succeeding point in a series.
d. Retains the actual units of measurement of the absolute data.
e. Reveals whether or not the data follow a consistent relative-change program.

Disadvantages are also mentioned in the reference and the reader is advised to note them. Most of them deal with user function, but training generally overcomes this stumbling block. In noting the advantages of the ratio or semi-logarithmic chart, it is evident that any unit distance (i.e., one-half inch vertical measure) represents the same amount of change any place on the chart. This can be confirmed by simply measuring one-half inch vertical distance from .01 to .04 and then measuring one-half inch vertical distance from 5.0 to 20.0. The equal distance shows equal change: .01 x 4 = .04; 5.0 x 4 = 20.0. Other measures would also illustrate this partitional relationship.

Summary and Application

With the possession of a behavioral ruler (movement cycles) and a chart for visually displaying our measures (six-cycle semi-logarithmic paper), we are prepared to view the learning process. The movements selected for analysis were measured and charted at the Experimental Education Unit of the University of Washington by members of the teaching staff and interested students.

Figure 7 illustrates the monitoring of a 19-year-old boy's movement cycles in writing letters. The boy, recovering from the effects of an accident in which he had suffered severe head injuries, worked at his writing exercises approximately 25 minutes a day for 19 weeks. Lyons and Carchan's Write and See workbook and a chemical pen were used. The work was presented under the direction of Mr. William Hulten and Mrs. Ann Mingo, and movement cycles were recorded on the chart. The environmental conditions were so arranged that no specified consequence followed any movement cycle other than the boy's proceeding to the next tracing position.

The data clearly indicate that the correct movements per day increased (shown on the chart as median movements per week). The errors, (tracings that produced a yellow mark) decreased as the 19-week period progressed.

Figures 8a-8d depict movement cycles related to walking. Some children, because of handicaps, walk awkwardly and this is often a source of social embarrassment. To help such a 16-year-old boy, the teacher, Mr. William Hulten, consulted with Dr. Barbara Milacek of the University of Washington Physical Education Department, who suggested some movement cycles to improve the boy's walking. These included: placing the heel down first; placing left arm and right leg forward simultaneously (and the reverse); locking weighted leg and placing feet in a on-foot path. Figure 8a shows correct and error movement cycles of locking the leg: phase one, 3 weeks correct and error movements interspersed; phase two, instruction was given, error dropped, correct movements stabilizes; and phase three, after instruction, errors mostly at zero.

Figure 8b, placing feet in a one-foot path, shows change only in phase two during the instructional period; however, instruction did not maintain the performance of walking in a straight line. Figure 8c, placing the heel down first, reflected no performance change. Figure 8d shows a change in correct and error movements when placing opposite hand and foot forward, and the change was maintained after two day's instruction. The boy's walking has improved. However, at least two movement cycles remained unaffected by instruction.

Thus far, movements of "write" and "place" have been shown on the charts. Figure 9 shows the movement of "say." The movement cycle was to say phrases (of any length) without being told to do so by the teacher aide. The 10-year-old boy has a very low rate of verbal emission. The

Figure 8A

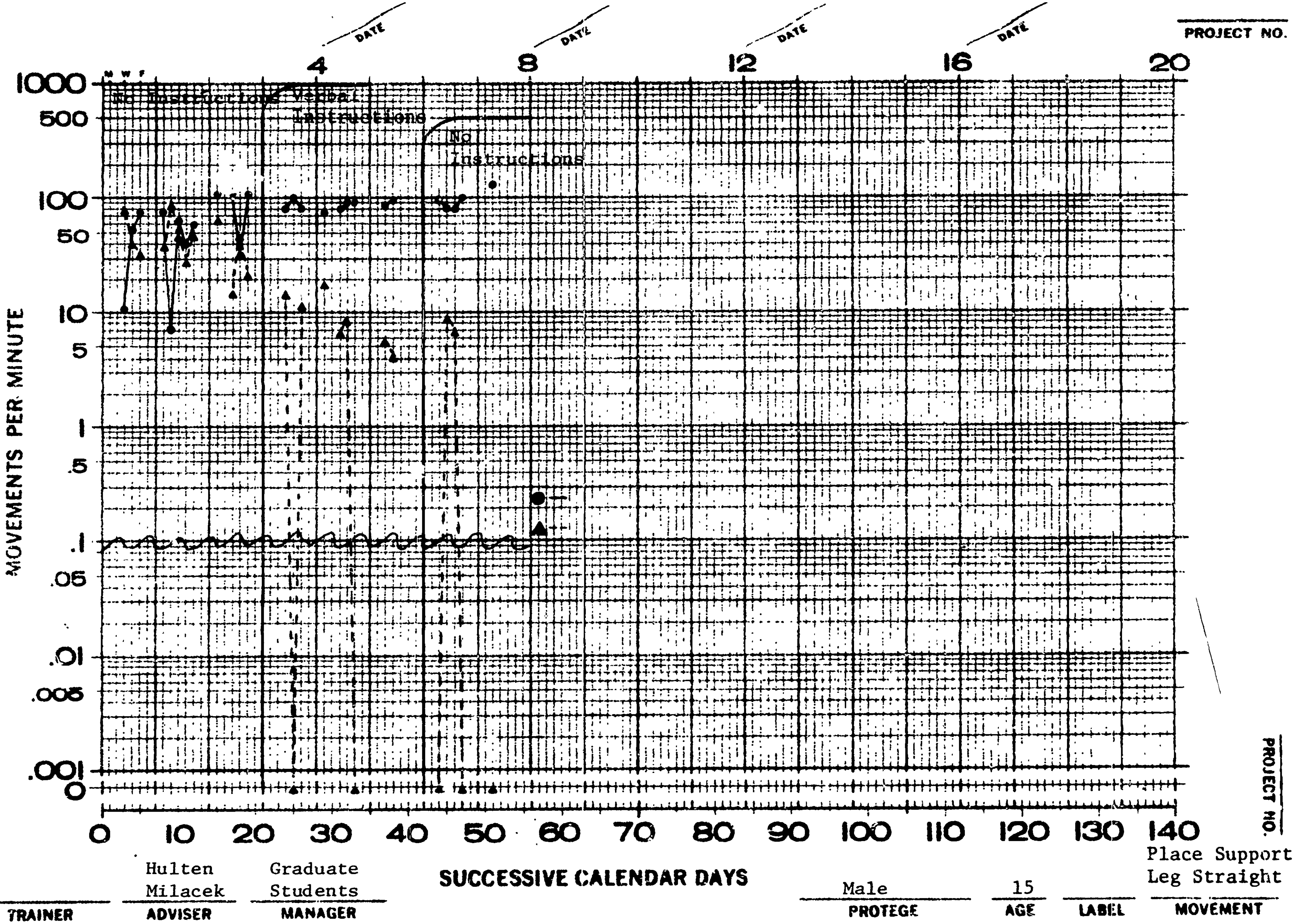
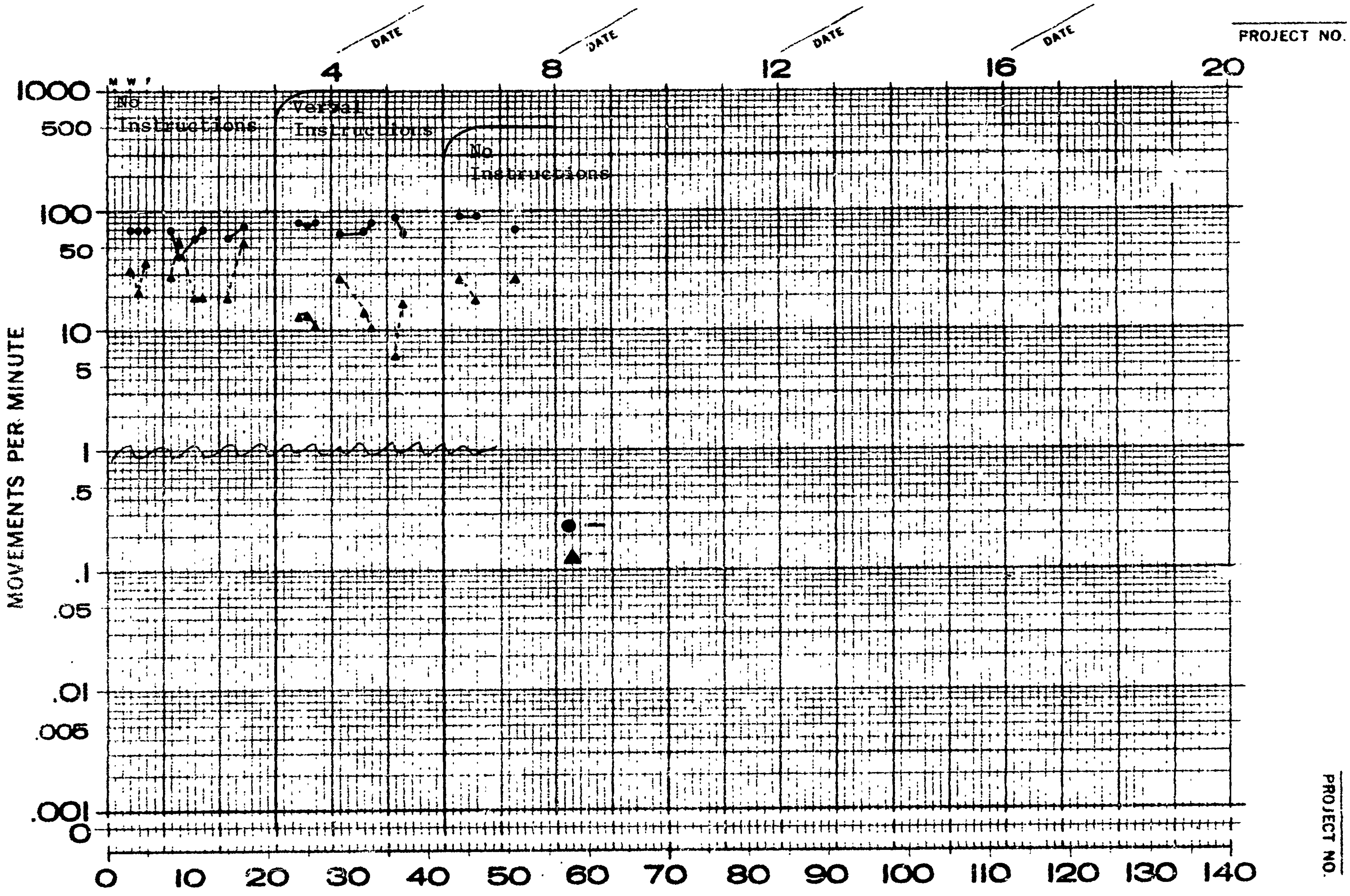
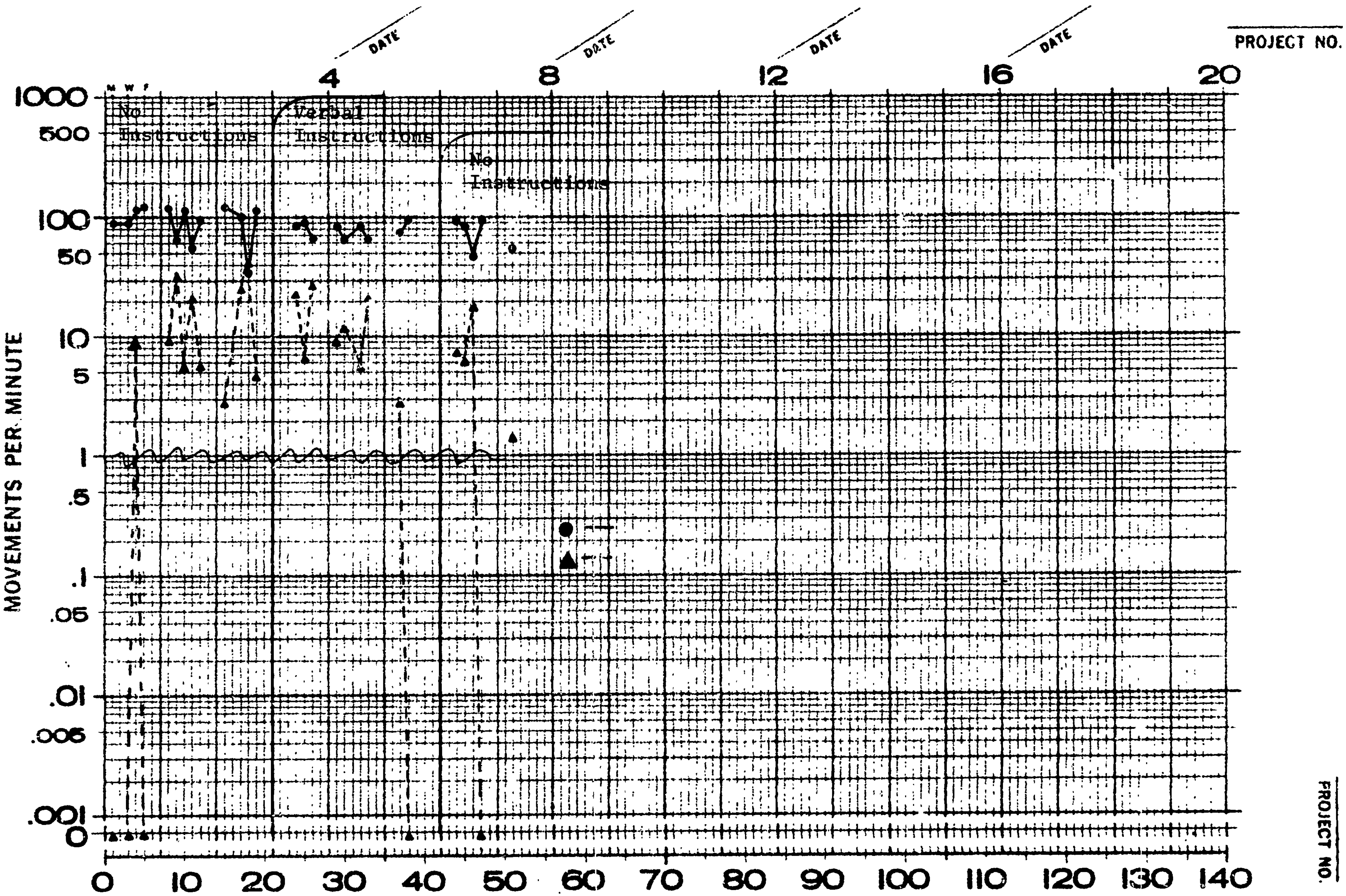

PROJECT NO.
DATE
DATE
DATE
DATE
4
8
12
16
20
M W F
No Instructions
Verbal Instructions
No Instructions
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
PROJECT NO.
Hulten
Milacek
Graduate Students
Male
15
Place Supporting Leg Straight
TRAINER
ADVISER
MANAGER
PROTEGE
AGE
LABEL
MOVEMENT

Figure 8B



TRAINER	ADVISER	MANAGER	PROTEGE	AGE	LABEL	MOVEMENT
	Hulten Milacek	Graduate Students	Male	15		Place Feet In One Footpath

Figure 8C



TRAINER	Hulten Milacek	Graduate Students	Male	15		Place Heel Down First
	ADVISER	MANAGER	PROTEGE	AGE	LABEL	MOVEMENT

Figure 8D

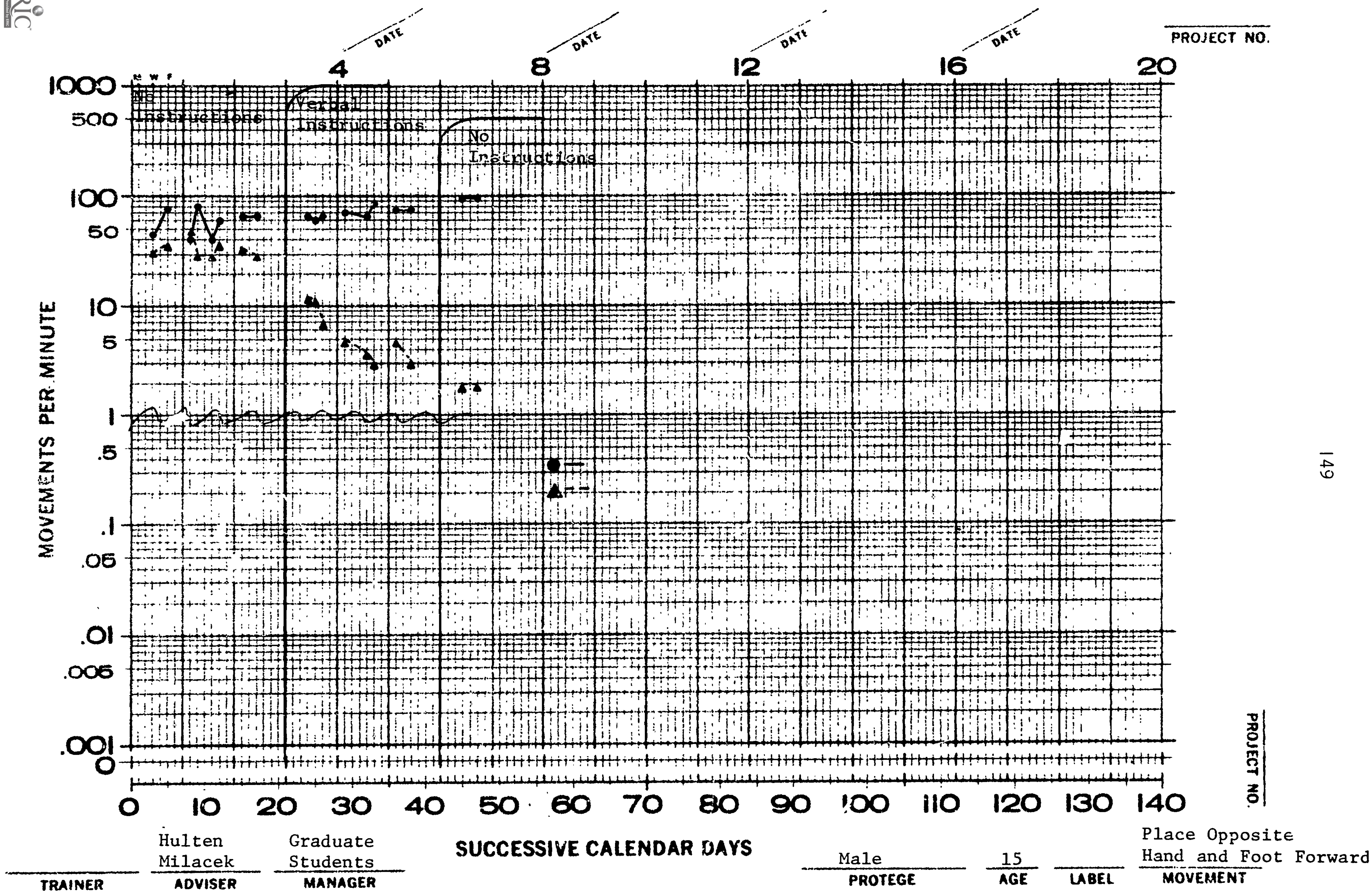
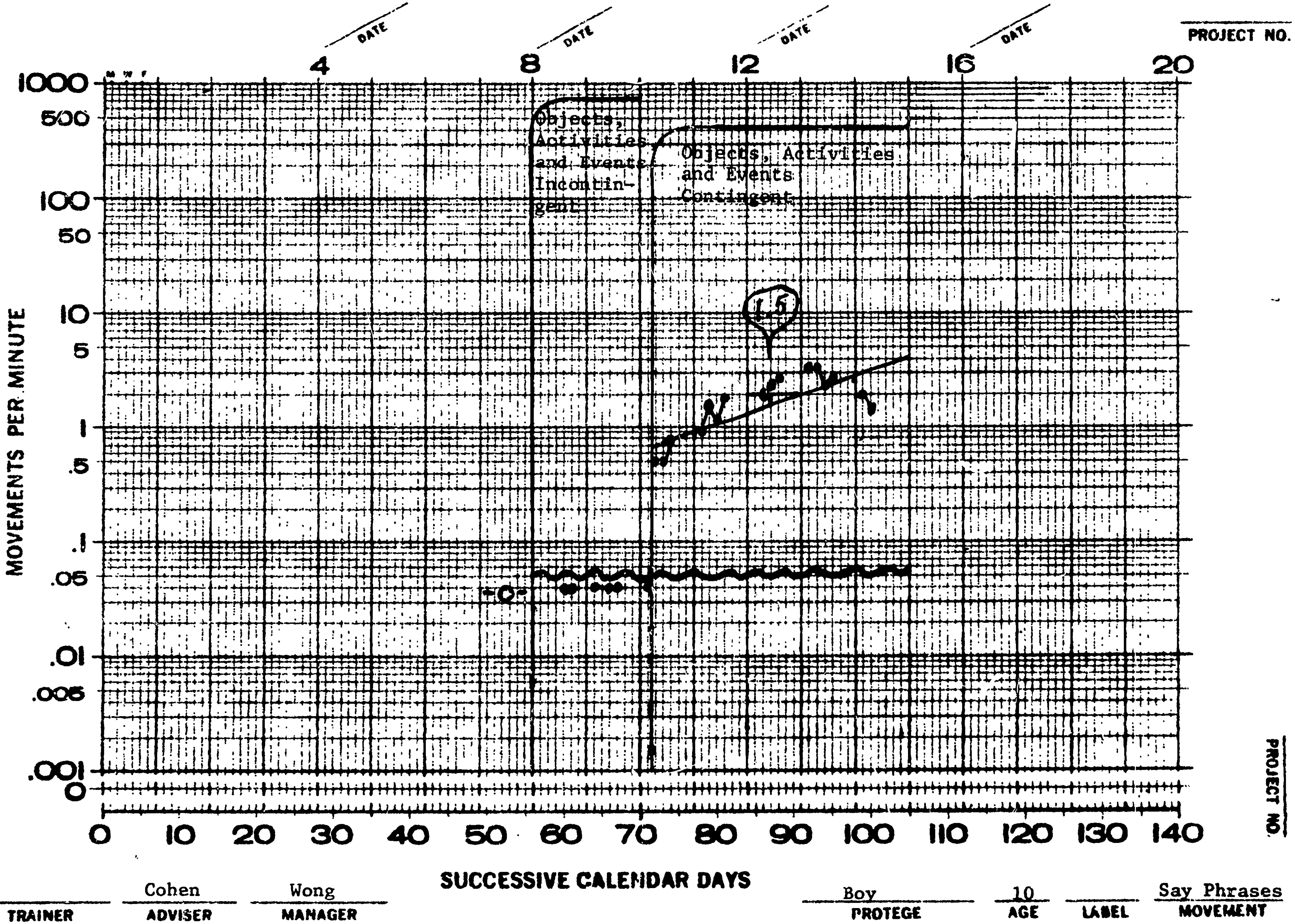

DATE
DATE
DATE
DATE
PROJECT NO.
4
8
12
16
20
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
No Instructions
Verbal Instructions
No Instructions
MOVEMENTS PER MINUTE
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
PROJECT NO.
TRAINER
Hulten Milacek
ADVISER
Graduate Students
MANAGER
Male
PROTEGE
15
AGE
LABEL
Place Opposite Hand and Foot Forward
MOVEMENT

Figure 9



PROJECT NO.
DATE
DATE
DATE
DATE
4
8
12
16
20
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
Objects, Activities and Events Incontingent
Objects, Activities and Events Contingent
1.5
-0-
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
PROJECT NO.
TRAINER
Cohen
ADVISER
Wong
MANAGER
Boy
PROTEGE
10
AGE
LABEL
Say Phrases
MOVEMENT

project ran about 20 minutes a day during the child's time off, which he had earned. As the chart shows, a zero "say" phase was recorded during the first six observation days. The zeros are plotted beneath the record floor line in phase one because the zero point is in essence non-existent in a ratio scale. The change in the environment for the child came when Mrs. Marilyn Cohen, the teacher, placed any event, object or activity on a contingency basis. If the child said what he wanted, the teacher aide was to permit the event, object, or activity to occur. Phase two shows an excellent change. The boy was asking for things at a median rate of 1.5 movements per minute. Although the teacher was pleased that a technique for increasing the child's verbal behavior worked, Mrs. Cohen was faced with a data picture in phase two that was disconcerting. The boy, while accelerating his verbal performance, did it at a decreasing rate. The "change index" shown by the triangle formed between the center of the phase line, the median of the phase line, and the trend of the phase line shows less and less acceleration. Possibly another change is needed. An examination of the learning process charted on six-cycle log paper reveals information that makes decisions inevitable, by the child, by the teacher, or by both of them.

As educators, we can no longer discuss individualized instruction with references to group norms. We can no longer build environments on group standards when we know how to tailor environments for individuals within the group. And further, as educators, we have a clear choice. That is, to use the information provided by a behavioral ruler or to ignore its existence.

References

American Society of Mechanical Engineers. American standard time-series charts (ASA Y15.2). New York, 1960.

Haughton, E. Personal communication, 1968-69.

Kunzelmann, H. P. A strategy for experimental teaching in an institutional setting. The Journal of Correctional Education, 1967, 19 (4).

Lindsley, O. R. Direct measurement and prosthesis of retarded behavior. Journal of Education, 1964, 147, 62-81.

Skinner, B. F. Science and human behavior. New York: Macmillan, 1953.

Smith, F. M. and Adams, S. Educational measurement for the classroom teacher. New York: Harper & Row, 1966.

CHAPTER VII INTRODUCTION OF THE BEHAVIOR MODIFICATION APPROACH TO SPECIAL EDUCATION: A SHAPING PROCEDURE

Frank M. Hewett

Special education as an emerging discipline over the past century and a half has been dedicated to "making a difference" in the lives of children with physical, sensory, intellectual, or behavioral deviations, and today reasonable optimism exists that all such children can profit from an educational program. Such determination and optimism have developed amidst a humanistic tradition which has aimed at respecting and protecting the exceptional child as an individual and viewing his problems in a broad physical, social, emotional, and intellectual context. As a result, disciplines of pediatrics, sociology, psychology, psychiatry, and neurology have been called on by the special educator for assistance in diagnosis and assessment as well as in the formulation of educational goals and the development of methodology. While the assistance of these disciplines has been invaluable in many instances, some teachers of exceptional children have found themselves in the middle of a complex, multi-disciplinary maze, the choice points of which are labeled in such alien jargon and which involve such sophistication with extra-educational concepts that an efficient, understandable route to the ultimate goal (e.g., a useful and practical educational program for the child) is most difficult to arrive at.

Recently, an approach to the education of exceptional children, focusing on their observable behavior as it is manifested in the classroom rather than antecedent psychological, and social causal factors, or measurable or inferred physical or neurological deficits, has become of interest to special educators, especially in programs for the mentally retarded and the behaviorally disordered (Birnbrauer, J., Bijou, S., Wolf, M. and Kidder, J., 1965; Cohen, H., Goldiamond, I., Filipczak, J. and Pooley, R., 1968; Haring, N. and Lovitt, T., 1967; Hewett, F., 1968; Lent, J., 1966; Patterson, G., 1965; Quay, H., 1966; Risley, T., 1968; Whelan, R., 1966). This approach is based on knowledge regarding the science of learning acquired in the experimental psychological laboratory (Skinner, 1953). It is direct in its implications for teaching children and changing their behavior: 1) select a terminal goal (e.g., reducing out-of-seat behavior, or improving self care or academic skills), 2) prepare a series of tasks involving reasonable increments which lead up to such a goal, and 3) through careful selection and presentation of stimuli and consequences, modify the child's behavior and bring it in line with that goal. The directness of this behavior modification approach is not matched by simplicity, for it also involves sophistication in concepts and familiarity with terms not usually found in the background of the special educator. In addition, its technological focus shifts emphasis from holistic concern with the child and preservation of a humanistic tradition. This shift in emphasis, however, by no means precludes these latter considerations. What is involved is a shift in point of view. The special educator may have difficulty in adopting such a point of view due to numerous reservations the field of education appears to hold regarding a behavioristic approach: 1) children are not animals and learning theory has largely emerged from the animal laboratory, not the human classroom, 2) response measurement and manipulation of stimuli and consequences in the classroom dehumanizes the teacher and provides a "technician" rather than "teaching artist" role, 3) if direct behavior modification goals dominate an educational program, the child's individuality, and opportunities for self expression and creativity are apt to be neglected, 4) emphasis on a powerful technology in education may be destructive if irresponsibly utilized, 5) rewarding children constitutes "bribery" and bringing about behavior change through manipulation of environmental events and inducements is "brainwashing."

Discussion of each of these reservations is subject matter for a separate paper or papers and will not be attempted here. Suffice to say that proponents of the behavior modification approach have their work cut out for them as they attempt to bolster the efforts of special education to "make a difference" in the teaching of children with learning and behavior problems, particularly in the public school.

Public school awareness and acceptance of research-based knowledge and theory in learning and behavior has been both puzzling and disappointing over the years. However, dissemination of such knowledge and theory in remote journals, use of purist experimental and theoretical terminology, and "holier than thou" attitudes on the part of "ivory tower inquirers" in relation to problems of public school application and service have not aided the situation. Just how difficult it is to bridge this gap in knowledge and theory is illustrated by this writer's observation that fundamental statements regarding individual differences which usually appear in the opening pages of most basic psychology texts have not been utilized to make the difference they could in school programs for all children. A related problem centers around education's preoccupation with "ideas" and "lofty goals" rather than efficient technology. Charters (1945, 1948) in the 1940's described the "idea men" and the "engineers" in education and commented on the fact that the former assume that if ideas and goals are formulated someone else will put them into practice and that if they give the field a new idea, a sufficient contribution has been made. The latter display a systematic and "patient thoroughness" and objectivity regarding the evaluation of such goals and a concern with their attainment. Skinner (1968) is a modern day critic of the lack of balanced emphasis on ideas and goals and effective means of achieving them in education.

Perhaps one line of attack with regard to gaining awareness and acceptance of the behavior modification point of view and exploitation of the powerful methodology it offers is to approach the field of special education and the teachers in it in the same manner one might approach a child whose behavior we sought to modify. The terminal goal might be acceptance, understanding, and use of behavior modification in the education of children with learning problems. Determination of a series of steps to take in order to achieve this goal would involve starting close to where the field and the teachers are now and gradually moving them toward more sophisticated understanding and use of the approach. Finally, manipulation of consequences to accomplish this goal may not be as complicated as it might seem. Introduction of an approach for increasing the probability of successfully making a difference in the education of children with learning problems into a field frustrated with its failures and confused regarding improving its effectiveness will undoubtedly provide a powerful positive reinforcement.

The incorporation of a behavior modification orientation in special education is desirable for several reasons. First, it assigns the teacher the role of "learning specialist," far closer to the role teachers are prepared for and are expected to fulfill from that of "junior psychiatrist" or "pseudoneurologist." Second, it resists preoccupation with asking questions regarding why the child has a learning problem and directs the teacher to get down to the business of providing an answer to, "How can I increase the child's ability to adapt successfully?" Third, it eliminates reliance on ominous medical, psychiatric, and neurological labels which simply have no translatability into educational practice and which imply the learning problem is primarily the child's not the teacher's responsibility. Thus, a behavior modification point of view holds promise for promoting an increased learning emphasis in the classroom. Where to start in introducing it to the teacher and the field of special education?

An obvious starting point is the introduction of the basic terms (e.g., stimulus, response, contingency, and reinforcement) and processes (e.g., successive approximation, shaping, response measurement, scheduling of reinforcement, and extinction) and have these related to teaching and learning in the classroom. When this occurs early in the university or college teacher preparation period, it may be extremely useful, but when it occurs at the close of the teacher preparation period or years after a given teacher has begin teaching in the public school, its usefulness is debatable. Teaching styles and biases form fairly rapidly and acceptance of new approaches may meet with considerable resistance. It may be that certain critical aspects of the behavior modification approach can be translated and emphasized so that they are understandable and acceptable to individuals already teaching in the field of special education, although more direct introduction of the approach into the early teacher preparation period certainly needs to be investigated. What is discussed in this paper is primarily a means of shaping the behavior of those already functioning in the field.

A major consideration in this shaping procedure appears to this writer to be an emphasis on goals toward which behavior modification might take place, rather than preoccupation with technology or methodology for accomplishing such goals. The guidelines offered by the behavior modifier (Ullman and Krasner, 1965) emphasize the importance of 1) identifying maladaptive behavior, 2) identifying environmental events which are supporting such behavior, and 3) modifying the environment so that more adaptive behavior is learned. "Identifying maladaptive behavior" is a big order for the average teacher. What are those behaviors which truly interfere with successful learning with all children? When this writer has asked this of large groups of special education teachers, he has obtained a lengthy list dominated by non-conforming, uncooperative behaviors and including a full description of problems seen in the classroom from distractability to academic deficits. All of these keep children from learning and teachers from teaching. Selection of certain target behaviors from such a list for modification should be relatively simple, but in this writer's experience priority rankings are difficult to establish and nebulous goals such as "improving citizenship" or "building reading comprehension" may be set more readily than specific behavioral goals such as "paying attention" and "following directions."

In an effort to provide teachers with a behavioral framework within which to view children which includes a statement of categories of behaviors which are essential to learning success, a developmental sequence of educational goals has been formulated as the first step in a shaping procedure to orient teachers to the behavior modification point of view. This sequence has evolved over an eight-year period of studying and teaching exceptional children, particularly the severely emotionally disturbed at the Neuropsychiatric Institute School in the Neuropsychiatric Institute of the University of California, Los Angeles. It is the result of contributions of both psychiatric and educational staff members and was primarily devised by this writer in collaboration with teachers in the Neuropsychiatric Institute School itself (Hewett, 1964, 1967, 1968). The sequence states that in ascending order of importance, children must acquire behaviors in areas of 1) attention, 2) response, 3) order, 4) exploration (thorough and accurate multi-sensory experiences with the environment), 5) social (behaviors which gain approval and avoid disapproval), and 6) mastery (cognitive and academic skills). Each of these areas or behavioral categories is fairly broad and includes many specific behaviors. For example, paying attention involves visual, auditory, or tactual orientation to the task and consideration of acuity, perception, and retention. The major contribution of the developmental sequence appears in organizing a behavioral framework within which the teacher may select target behaviors in a systematic and efficient manner. It is not viewed as an ambitious theoretical statement, although in its formulation, the positions of individuals such as Freud, Erikson, and Piaget were considered. While Bijou (1968) and Zigler (1954) have spoken out against pursuing the understanding of human development within the realm of hypothetical constructs and "grand designs," during the initial stage of moving educators toward a more empirical base, provision of some type of statement of goals, tentative as it may have to be, appears to this writer a crucial consideration. Rather than a compromise for the sake of expediency which can only delay application of scientific knowledge, this may be an essential first step in bridging the communication gap which exists between the "idea men" and "engineers" in education.

A second major concern in this shaping process may be expanding of the educator's appreciation of the importance of respondant as well as operant behavior in learning. Behavior modification as it is emerging in special education is primarily concerned with operant behavior. This type of behavior is observable, voluntary on the part of the individual, and subject to control by the consequences which follow it. Operant behavior makes up the bulk of the behaviors seen in the classroom. Respondant behavior is largely unobservable, involuntary, and conditioned according to Pavlovian principles. Respondant behaviors include the "feelings" elicited by certain stimuli in the environment.

Eight-year-old Henry, who is a total non-reader and who has been subjected to two previous years of frustration, devaluation, and unsuccessful attempts to teach him to read, is likely to have been conditioned in the respondant sense to become anxious, fearful, and possibly negative when asked to read. If on the first day of a new class, he is handed a third grade book and asked to read in front of a group by the teacher for purposes of assessing his reading level, we may expect him to become upset. If he breaks down and cries, struggles in vain with a word here and there while beads of perspiration form on his brow and his hands tremble, or angrily throws down the book and runs from the room, the effects of his previous conditioning will be highly visible. Skinner (1968) encourages concern with the operant behaviors which such a child exhibits (e.g., struggling with the words, throwing the book down, and running away) rather than attempting to explain such behaviors on the basis of respondant conditioning. This will, in effect, focus the teacher on the aspects of Henry's behavior which can be seen and about which something can be done in the classroom and make apparent that by manipulation of stimulus events (e.g., recognizing Henry's reading problem and providing him with a pre-academic training program) and consequence events (e.g., guaranteeing Henry his full share of positive reinforcement for success at whatever level of beginning reading instruction he can handle) his behavior in the classroom can be altered.

However, the notion that certain conditioned stimuli such as teachers, books, assignments, and grades do elicit respondant behavior in children with learning problems may be particularly helpful to emphasize as the behavior modification approach is introduced into special education. The technique of desensitization explored by Wolpe (1965) and Eysenck (1960) and earlier by Jones (1924) involves the establishment of behavior which is totally incompatible with respondant behaviors such as fear and anxiety. If, following exposure to the level of the conditioned stimuli, the individual can tolerate with a minimum of discomfort (e.g., perhaps coming into the classroom in Henry's case) a response which is incompatible with the discomfort experienced in relation to the conditioned stimuli is engaged in (e.g., Henry's participation in an appealing arts and crafts project while other children are reading), a beginning step in the deconditioning process is underway.

This writer has worked for three years with the staff of the Santa Monica Unified School District in California in the development of an engineered class design to both decondition educationally handicapped children to stimuli associated with school and previous failure and conditioned essential learning behaviors as described by the developmental sequence of educational goals (Hewett, Taylor and Artuso, 1967). Throughout the day consequences in the form of check marks (exchangeable initially for tangible rewards, later for privilege time, and finally for a report card) are given every fifteen minutes. As a given child exhibits distress or frustration with a particular assignment, the teacher follows a systematic series of interventions which involve changing the assigned task until the child exhibits more appropriate behavior. If he can successfully be channeled into an alternate activity, he receives his full complement of check marks with no penalty for his failure to complete the initial task assigned. Such a procedure runs the risk of reinforcing maladaptive behavior in terms of operant conditioning since the consequences for balking at the original assignment may result in rewarding consequences. But viewed in the context of the respondant model, the change in the stimulus situation is directed toward maintaining a response on the child's part which is incompatible with the previously established respondant of fear and anxiety associated with school.

A lengthy discussion of just where the respondant behavior leaves off and the operant behavior begins in such a situation is seen by this writer as academic. The field of special education, particularly with the emotionally disturbed, has been dominated by psychodynamic psychology for several decades. Whether behaviorists like it or not, how the child "feels" about himself, others, and learning is not likely to suddenly disappear from the list of major concerns of the special educator. The behavior modification approach recognizes that such feelings do indeed exist and can lawfully explain their development by means of the respondant or classical conditioning model. Such a lawful explanation, therefore, can be utilized to bridge still another gap and facilitate the shaping procedure discussed in this paper. In addition, the desensitization process may be extremely useful in setting the stage for more effective operant conditioning and learning. In this writer's experience, few children ever become preoccupied with "beating the system" in the engineered classroom because it "pays off" when you don't do an assignment. When they do, use of negative consequences such as a time-out period during which no check marks are given may have to be used, but this has been done very seldom. The real value of the intervention procedures used seem to center around the message of the teacher and the school to the child, "We will not let you fail.' This message appears to far overshadow the message, "Not working pays off." The differential effect of these two messages probably has something to do with the difference between children on the one hand and rats and pigeons on the other.

A third and final consideration in the shaping procedure for introducing the behavior modification approach into special education is concerned with translation of the basic principles of the approach into specific classroom programs and practices rather than reliance on trial and error application of the principles following a broad indoctrination of the teacher. Such specificity has its counter part in teaching machine programs and programmed instructional units devised according to learning theory principles and presented to the teachers in "package" form for classroom use.

The engineered classroom design developed in the Santa Monica schools for educationally handicapped children is an example of a "package" approach. Teachers who teach in engineered classrooms are trained to assess the children according to the developmental sequence discussed earlier and to use a simple but largely familiar task-reward-structure methodological frameword in the application of behavior modification principles. No attempt is made to make such teachers "operant conditioners" or "experimental psychologists." Rather, efforts are directed toward helping them make a difference with the children in their classrooms through step by step procedures including classroom floorplan, curriculum, schedule, management techniques, and a check mark system in lieu of traditional grading practices. It has been observed that perhaps such an approach actually does as much if not more for the teacher than the child. While the child is exposed to a highly predictable and individualized learning environment in which terminal goals are worked towards in successive approximations and consequences clearly linked to task accomplishment, the teacher is aided with a plan designed to take full advantage of her resources and expertise and to minimize problems of task selection and control. Such a plan not only clearly organizes the teacher's effort but greatly increases the probability that these efforts will be successful and personally rewarding.

No such tailoring of teacher to design appears to accomplish the much feared dehumanization some educators worry about. The design is primarily useful in "launching" both teacher and child into successful teaching and learning. This writer has never seen a behavior modification program in education that ultimately did not have a great deal to do with the individual personality of the teacher rather than only the administration of certain principles and procedures. Eventually the "package" is utilized by the teacher in a unique and individualized manner and in one sense the ultimate contribution of behavior modification to education may be to help special educators be themselves more efficiently and effectively.

This paper has briefly attempted a discussion of the introduction of the behavior modification approach to special education. The humanistic, multi-disciplinary heritage of the field makes direct implementation of an empirical approach difficult to accomplish overnight, and greater emphasis on the delineation of goals, on unobservables or feelings in children with learning problems as being largely conditioned, and specific development of classroom designs to introduce both the teacher and the child to the principles in action have been suggested as shaping procedures which may increase the probability that the behavior modification approach will achieve a position of usefulness in special education.

REFERENCES

Bijou, S., Ages, stages and the naturalization of human development. American Psychologist, 1968, 23, 419-428.

Birnbrauer, J.S., Bijou, S.W., Wolf, M. M., and Kidder, J. D. Programmed instruction in the classroom. In L. P. Ullman and L. Krasner (eds.), Case studies in behavior modification. New York: Holt, Rinehart, and Winston, 1965.

Charters, W. Is there a field of educational engineering? The Educational Research Bulletin, 1945, 24, 29-37.

Charters, W. Idea men and engineers in education. Educational Forum, 1948, 12, 399-406.

Cohen, H., Goldiamond, I., Filipczak, J., and Pooley, R. Training professionals in procedures for the establishment of educational environments. Silver Spring, Maryland: Educational Facility Press, 1968.

Eysenck, H. (ed.) Behavior therapy and the neuroses. New York: Macmillan, 1960.

Haring, N. and Lovitt, T. Operant methodology and educational technology in special education. In N. Haring and R. Schiefelbusch (eds.), Methods in special education, New York: McGraw-Hill, 1967.

Hewett, F. A hierarchy of educational tasks for children with learning disorders. Exceptional Children, 1964, 31, 207-214.

Hewett, F., Taylor, F., and Artuso, A. The Santa Monica project: demonstration and evaluation of an engineered classroom design for emotionally disturbed children in the public school, Phase I: Elementary level. Final Report: Project No. 62893, Demonstration Grant No. OEG-4-7-0628930377, Office of Education, Bureau of Research, U. S. Department of Health, Education and Welfare, 1967.

Hewett, F. Educational engineering with emotionally disturbed children. Exceptional Children, 1967, 33, 459-467.

Hewett, F. The emotionally disturbed child in the classroom: A development I strategy for educating children with maladaptive behavior. Boston: Allyn and Bacon, Inc., 1968.

Jones, M. C. A laboratory study of fear: The case of Peter. Pediatrics Seminar, 1924, 31, 308-315.

Lent, J. A demonstration program for intensive training of institutionalized mentally retarded girls. Progress Report, Grant MR-I 801 A66, Parsons State Hospital and Training Center, Parsons, Kansas, 1966.

Patterson, G. An application of conditioning techniques to the control of a hyperactive child. In L. Ullman and L. Krasner (eds.), Case studies in behavior modification. New York: Holt, Rinehart, and Winston, 1965.

Quay, H. Remediation of the conduct problem child in the special class setting. Exceptional Children, 1966, 32, 509-515.

Risley, T. Learning and lollipops. Psychology Today, January, 1968.

Skinner, B. Science and human behavior. New York: Macmillan, 1953.

Skinner, B. The technology of teaching. New York: Appleton-Century-Crofts, 1968.

Ullmann, L., and Krasner, L. (eds.) Case studies in behavior modification. New York: Holt, Rinehart, and Winston, 1965.

Whelan, R. J. The relevance of behavior modification procedures for teachers of emotionally disturbed child. In Knoblock, P. (ed.), Intervention approaches in educating emotionally disturbed children, Syracuse: Syracuse University Press, 1966.

Wolpe, J., Salter, A., and Renya, L. (eds.) The conditioning therapies. New York: Holt, Rinehart, and Winston, 1954.

Zigler, E. Metatheoretical issues in developmental psychology. In Mark, M. (ed.), Theories in contemporary psychology. New York: Macmillan, 1963.

CHAPTER VIII EFFECTIVE TEACHING OF CHILDREN WITH BEHAVIOR DISORDERS

R. J. Whelan

It may be asserted that teaching involves only the conveying of information which others are expected to learn and practice in subsequent situations. If teaching could be reduced to such a simple process, it would be relatively simple to evaluate the effectiveness or efficiency of the procedure. Information could be conveyed verbally and motorically to an individual who is deficient in a particular skill. The acquisition or learning of the information could be ascertained if the individual recites the material with a reasonable degree of accuracy. Evaluation of the process may be implemented by direct observation of the individual in situations where the newly acquired skill should be used. If the individual displays the skill at the appropriate time, place, and circumstance, the teaching process may be judged successful. However, if the individual remains deficient in utilizing the skill, the teaching process should be judged unsuccessful.

Unfortunately, the acquisition of knowledge and skill is often dependent upon the processes described. There is the exception, though, in that the evaluation aspect of the process is either haphazardly instituted, or not attempted. Teaching does subsume the processes described, but it also includes more. Teaching must be concerned with the information conveyer (teacher), the information recipient (learner), and the environment in which interaction or transaction occurs. As such, interaction among the teacher, learner, and the environment is a continuous process even though it may not be recognized or preplanned.

Prepared with Patricia A. Gallagher

The major focus of subsequent sections is the delineation of procedures which can be used to systematically manage the interaction and transaction that occur among teachers, children, and the environment, the three main components of the learning process. More specifically, the purposes and foci are as follows:

1. Present information which will convey the importance of understanding the principles and techniques used in observing behavior.
2. Describe procedures for systematic application of behavior principles to behaviors which need to be changed.
3. Discuss methods which can be used to evaluate the effects of implementing behavior principles and procedures to change behavior.
4. Describe methodological aspects which teachers may implement in enhancing learning by managing teaching, child, environment interaction.

The Educational Challenge

School and its representatives, teachers, are often the first community agents to establish contact with children. When children leave home and spend a significant portion of the day in school, they are expected to accomplish three objectives which are as follows:

1. Leave the protective milieu of the home and venture into relatively unknown, uncertain areas of expectations for behavioral performances.
2. Establish appropriate interpersonal relationships with peers, and adults in various authority figure roles.
3. Acquire and accumulate skills and knowledge which will culminate in behaviors that are necessary for becoming a contributing member of society.

Children are expected to change from various stages of dependent, unsocialized, self-centered behavior to points upon the behavior continuum which represent degrees of independent, socialized, and group-centered behavior (Whelan, 1966b). Meeting these objectives requires investment and commitment from children and teachers. The business of children is school in which learning and growth must progressively occur. The business of teachers is to prepare the classroom environment in such a manner that when children interact with it, learning and growth represent the culminating results.

It must be recognized that teaching is a continuous challenge which involves development, change, and evaluation. As such, a finalization of the teaching process is never reached. The educational challenge is concerned with continuous teaching improvement, and searching for procedures which will enhance teaching effectiveness. While the central theme of this chapter is relevant to more effective teaching of children with behavior disorders, it is hoped that most aspects will also apply to all children.

Instructional technology (Skinner, 1968) has and will continue to be co-existent with classroom interactions and transactions. Part of the educational challenge is concerned with more systematic application of known and existent technology in classrooms for children with behavior disorders. That is, how can teachers learn and apply existing knowledge in a manner which will promote effective teaching as measured by teacher and child continuous progress? While there is a serious lag between identified effective teaching procedures and their application in classrooms, a more serious problem is also apparent. A myriad of procedures can be described which will enhance learning, but such procedures are selected and applied by individuals. Awareness of procedures does not guarantee subsequent application, nor does it automatically assume that procedures will be applied correctly and systematically. Application of effective teaching procedures is a function of individuals who have the capability for error as well as accuracy.

Effective teaching involves understanding of several central concepts, the various behavior change procedures, and finally, correct application of concepts and procedures in classroom environments. When application occurs, concurrent evaluation is possible. Application and evaluation will result in further improvement of teaching, and should function to reduce the gap between discovery and implementation. That is, information pertaining to more effective teaching methodology could emanate from the same situation where it will be subsequently applied. Discovery of new procedures can and should take place in classrooms where child, teacher, environment interaction naturally occurs on a minute, hour, daily, and weekly basis.

The educational challenge can be clearly delineated. In view of the complexity of educational objectives, expectations for child performance in school related tasks, and selecting effective teaching procedures, it is not difficult to discern possible debilitating child, teacher, environment interactions which can affect children with behavior disorders. Educators must plan the environment to promote positive results from such interactions.

Defining Children with Behavior Disorders

It is important to be cognizant that definitions may serve as guidelines or stimuli which are thought to be functionally related to subsequent behaviors. Given an exact definition, it is reasonable to assume that this will prescribe certain procedures when interaction occurs with the object or person defined. Unfortunately, definitions are usually not that precise, nor have they promoted exact or effective behavior change procedural prescriptions.

In the continuing divergent discussions regarding treatment efficacy, the point is humorously conveyed by Eysenck (1961) who cites Raimy's definition of psychotherapy as "an unidentified technique applied to unspecified problems with unpredictable outcomes. For this technique we recommend rigorous training." While adherents of psychotherapy intervention in children with behavior disorders usually define patients more exactly, the quotation does emphasize the problems which can be encountered when excessive reliance is placed on definitions. In essence, definitions should serve to guide and assist individuals who are assigned to promote behavior change and more successful adjustment of children with behavior disorders.

Educationally relevant definitions of children with behavior disorders or emotional disturbance are provided by several sources (Bower, 1960; Haring and Phillips, 1962; Pate, 1963). Such definitions do have relevance, but do not stand independently in terms of automatically conveying behavior change procedures. That is, the definitions provide convenience in circumscribing an issue, but do not add specific knowledge relevant to methodology to resolve the issue. Further investigation beyond the definition is required to ascertain the various behavior change methods which can be used in interaction with subjects so defined. The definition does not prescribe the specific behavior change procedure. In fact, it is possible to obtain convergence about a definition, and then discover divergent descriptions of proposed effective methodology. The co-existence of different methods to solve similar problems should be recognized.

Problems with definitions. Definitions are not positive or negative in isolation from other variables. They are selected by individuals who may use them inappropriately or appropriately. An inappropriate use is exemplified by the unrecognized effects of the reification process. Reification is a three step process which is as follows:

1. A label or name is given to a single behavior or a list of behaviors.
2. The label may then be considered to represent the behavior or behaviors.
3. Eventually, the label is used to explain the occurrence of the behavior or behaviors.

A specific example will clarify the circularity of the reification process. Children are referred for diagnostic evaluations, and behavior change planning, to the Children's Rehabilitation Unit, University of Kansas Medical Center. One child may exhibit all of the classical behaviors associated with the label "autistic" (Kanner, 1943; Rimland, 1964). After diagnostic data are accumulated, a diagnosis of autism may be confirmed. Step one of the reification process has been completed. At this reference point, the use of the label or definition does not have deleterious effects. However, step two and three can exist unless awareness of possible misuse of the label is ensured. The diagnosed child may be placed in a preschool intervention program which is designed to foster more adaptive behavior in the school and home. Since the school program is experimental as well as service and training oriented, various professional visitors request to observe the program. Reification occurs and the cycle is completed when the visitor is told that the behavior being observed is caused because the child is "autistic". The label has been used to explain the behavior; step two and three have been completed.

Labels or definitions should be used as brief communication devices. A single word may be used to describe the presence of a syndrome or list of behaviors. However, agreement should exist among those individuals who select to use a label as to what specific behaviors are included within the rubric of the label. If a visitor uses the label "autistic," it behooves others to ask what behaviors must be exhibited before the label is assigned. Essential disagreement may become apparent, and must be resolved before there is functional commonality in the use of a label. Labels are used appropriately to describe behavior, and used inappropriately to explain behavior.

Functional definitions. In the previous section it was noted that definitions should suggest or prescribe remediation and corrective behavior change procedures. Exact specificity for each problem encountered in each child is probably not possible at this time. If it were, parents, aides, teachers and others would only have to refer to an electronic type of "trouble-shooting manual" to find the specific "trouble" and then implement the "repair." All of this implies that the nature or current status of the individual child is not considered, and that the behavior change process is automatic. While this may be a comfortable state for individuals who seek simplicity in decisions regarding procedures for changing behavior, it is deceptively simple. It is possible to apply known behavior principles to a problem, but the principles merely serve as guidelines for determining specific behavior change techniques that are functionally related to the problem and the child who exhibits it. Behavior principles form the guidelines for behavior change procedures, but the exact techniques are applied on the basis of individual child needs. It is important within this context to realize that behavior principles are few in number, are easily understood and verbalized, but then application to children can be quite complex, especially in requirements for careful, systematic planning of child, teacher, environmental interactions. Subsequent sections will discuss these aspects in more detail.

Present knowledge does allow some specificity in connecting definitions to intervention procedures. Children with behavior disorders (social and academic) are often viewed as being emotionally disturbed. The term "emotional" has a plethora of meanings, but perhaps the most functional one is evaluations placed upon experiences (Phillips, Wiener, and Haring, 1960). Intraindividual and interindividual responses which are labeled "emotional" are probably responses to present situations, or situations which are similar to painful or pleasurable past encounters. A child who, completing a reading assignment successfully, receives recognition from parents and teacher will probably display emotions which are described as positive. If the child fails in reading and is punished or harshly reprimanded, behaviors of a negative nature may occur. The child may avoid future efforts or may attempt to escape from a reading lesson by getting out of his seat and verbalizing a physical complaint.

It is proposed that the terms "emotional" and "disturbance" provide a functional method of connecting, in a general manner, a definition with subsequent programs of behavior change. The E in "emotional" represents behavior excesses; D in "disturbance" represents behavior deficits. Children with behavior disorders are characterized by excesses and deficits. They display some behaviors which parents, teachers, and peers would like to reduce or delete, excesses. They do not display some behaviors, or do not have them in their repertoire at a level commensurate with expected behavioral performance, deficits.

The effective teacher of children with behavior disorders is proficient in identifying and selecting behavior excesses and deficits that interfere with children's progress in school. Excesses and deficits are circumscribed as exactly and discretely as possible. A brief case description will convey the importance of this aspect. A child in a special class was described as introverted and withdrawn. This may be considered a deficit in social behavior or social interaction. It was decided to obtain more exact information pertaining to the degree of withdrawn behavior. Through direct observation of the child, it was determined that over a period of 120 hours of observation (24 school days) only seven child initiated social contacts occurred (Gallagher, 1968). The degree of withdrawn behavior was systematically defined, observed, and recorded. The behavior deficit was precisely delineated, and the teacher possessed a very discrete and observable behavior which could be changed by effective application of behavior principles. It would have been extremely difficult for the teacher to change withdrawn behavior without the exactitude with which the words "introverted" and "withdrawn" were translated into quite observable behavior deficits. Knowing this information, the teacher instituted a behavior change program designed to increase the frequency of appropriate social contacts. The behavior change program was in effect for 22 days. During the last 11 days of the program the number of appropriate social contacts ranged from 5 to 10 a day. An after behavior change program check revealed that the child was maintaining 6 to 10 social contacts a day (Gallagher, 1968).

The withdrawn child, with the assistance of a behavior change program, developed social behaviors to a level that was acceptable. Social contact was no longer a deficit behavior. The behaviors were acquired during the behavior change program. During this behavior acquisition segment, natural environmental events which normally maintain an adequate frequency of social contacts were paired with the synthetic environmental events implemented as a part of the behavior change program. When the program was discontinued, natural environmental events were sufficient to maintain an adequate level of social contacts. The sequence of planning for this child included a functional definition and delineation of a behavior deficit, designing a behavioral change program to reduce the deficit, and determining if the behavior would be maintained after the program was terminated.

The brief description of a single case can serve to indicate the method by which a behavior definition, if described specifically, can be closely linked with the objectives and procedures implemented to change the defined behavior. Definitions and labels can be an important initial portion of the teacher, child, environmental interaction. However, they must be functionally related to subsequent planning of the interaction process.

Children with Excesses and Deficits

Before implementing or initiating behavior change programs for children with behavior disorders, it is necessary to determine the frequency and degree of observable behavior excesses and deficits. Conducting a survey of children and teachers is a procedure which may be used to obtain data pertaining to behavior disorders. Such a survey may include one child, or several hundred children. Teacher descriptions of children's excesses and deficits may be obtained in the same manner.

A total of 783 children, grades 1 through 12, were asked to list and describe specifically aspects of school which were of concern to them (Overshiner, 1968). From this group, 2,589 responses were recorded for analysis. The details of this study will be reported in subsequent publications; therefore, the results will only be summarized. Most of the children's responses, 56%, were concerned with school procedures. Included in these comments were items related to unfair or too many irrelevant rules, crowded conditions, and poor cafeteria food. Twenty per cent of the comments specified subject matter comments such as "too much work," and "the work is too hard or not explained by the teacher." The rest of the comments were related to complaints about unfair teachers, aggressive peers, peer group ostracism of other students, and self directed statements concerning inability to complete work on time. Each child included in the survey wrote very specific comments about excesses and deficits in the school program which are relevant to comprehensive planning for behavior change.

Teacher perceptions of child related problems also contribute valuable information that can be used in planning for more effective child, teacher, environment interaction. A partial list of teacher concerns for children's behavior excesses and deficits is reported in another source (Whelan, 1966b).

As one aspect of a mental health workshop (Whelan, Gallagher, Grosenick, and Kroth, 1968), 185 elementary teachers were asked to list specific child problems which interfered with productive school adjustment. As expected, boys represented 80% of the cases listed by teachers; girls accounted for 20% of the cases. Behaviors listed, based on teacher perceptions, which needed to be changed were assigned to academic and social categories. Of the total child behaviors which required alteration, 63% were social and 37% were academic. Examples given for social behaviors were talking out without permission, out of seat, touching, hitting, pushing, tripping, squirming, sucking thumb, low frequency of social interaction, and getting out of cafeteria line. Academic behaviors included incomplete work, late completion of work, poor accuracy, slow performance, sloppy writing, and not following directions.

Additional analysis of listed behaviors indicated that the teachers believed 57% should be decreased in occurrence or frequency. Of the 57%, social behaviors contributed 94%, while academic concerns represented only 6%. Teachers believed that 43% of the listed behaviors should be increased in occurrence or frequency. From the behaviors that required increasing, 77% were academic and 23% were social.

Teacher emphasis upon social behavior and reducing behaviors, in comparison to concerns pertaining to academic behaviors, and increasing behaviors, may be interpreted to mean that teachers are punitive, and only interested in arranging the classroom environment in such a way that children become only passive recipients of information. However, an alternative interpretation is probably more feasible and accurate. Teachers at the workshop were concerned with decreasing those behaviors which interfered with the primary assignment for children and teachers, learning. That is, listed social behaviors that needed to be decreased were of the type which interfered with the acquisition of competencies, skills, concepts, and knowledge. Teachers are responsible for arranging the classroom to provide maximum probabilities that learning will occur and be used in subsequent situations. Inadequate program organization, or child avoidance of interacting with potential learning transaction must be corrected and adequate child, teacher, environment interactions instituted.

Children and teachers can be specific in describing behavior excesses and deficits. However, the description is only the beginning of the behavior change process. The implication of exact behavior descriptions is that educators do not have to converse, or describe child behaviors, in vague or rather general terms. It is possible to circumscribe rather precisely and accurately those behaviors which interfere with the learning process.

The analysis of children and teacher perceptions of environmental and human interactions which interfere with learning may be summarized by listing two questions:

1. Are children with excesses and deficits behaving or reacting in a negative manner in order to avoid and escape from inappropriate school program planning and the presentation of unrealistic learning tasks or expectations for performance?

2. Have children with excesses and deficits learned to manipulate the learning environment in ways which lead to circumvention of some pain and delay of gratification inherent in learning or acquiring new skills?

If behavior (academic and social) expectations for children are not realistic in that the probability of successful completion is quite low, it is reasonable to assume that children will attempt to avoid or escape from such encounters. It is important, therefore, for the teachers to realize that learning objectives or competencies to be acquired by children must be arranged so that child, environment interactions will culminate in successful task completion and acquisition of skills. Children should be expected to complete those tasks which they are capable of accomplishing, and not those which will only lead to failure, defeat, or loss of self worth.

The important balance which teachers must achieve is between questions one and two. First, realistic expectations for achievement must be analyzed and implemented. If the child attempts to avoid interaction with expectations for performance, it may be due to aspects more directly applicable to question two. Learning may not be totally pleasurable, particularly when a new skill must be acquired. Exhibiting a newly acquired skill, however, may provide many pleasurable benefits. Teachers must be able to discriminate between behavior excesses and deficits which are elicited by (1) inadequate or unrealistic performance expectations, and (2) reluctance to engage in acquisition behaviors because it requires energy investment, uncertainty as to final results, and some time delay between requirements for task mastery and the benefits which result from task mastery. This discrimination can enable teachers to plan tasks more appropriately, and to recognize the difference between normal acquisition reluctance to engage in tasks, and gross avoidance behaviors which are evoked by unrealistic performance expectations.

Types of Children with Behavior Excesses and Deficits

Three types of children are usually encountered in relation to ascertaining behavior excesses and deficits. These children present problems which are directly related to the two questions previously discussed.

The first type of child is one who can succeed in task completion, but will not invest energy in completing or interacting with the task. Expectations for performance are realistic, but the avoidance behaviors associated with acquiring new skills have been successful in circumventing prior child, teacher transactions. The teacher, in this case, must search for environmental conditions which will operate as an incentive or motivation for task completion. For example, a ten year-old boy was given a standardized achievement test and scored at the third grade level. The teacher believed that this did not represent the child's capabilities. An alternate form of the test was given and the child was told that he could earn a penny for each correct answer. The score achieved was at the fifth grade level, an increase of two grades. It is not suggested that money be used in achievement testing situations. The example provided is only for purposes of illustrating the importance of arranging environmental conditions to maximize the probability of adequate task interaction and performance. Similar results, which will be presented in subsequent sections, can be obtained in the hourly or daily planning of programs for children with behavior disorders.

The second type is represented by those children who fail to complete tasks at an accuracy level which is adequate. However, they appear to be highly motivated in that observation of their performance does not reveal attempts to avoid tasks by rushing through the assignment and responding haphazardly. Observations further reveal that such children are expending energy and effort, but are unable to reach criterion for success. Adding synthetic or extrinsic incentive aspect, e.g., money, does not alter the overall performance score. These children appear to be making a reasonable effort, are intrinsically motivated to complete tasks accurately, but do not achieve. It is obvious that the program or task expectations must be altered. If a child does not have the prior skills or competencies necessary to complete a task, failure is inevitable. Asking a child to complete a long division problem when he does not understand concepts related to division or does not have competency in subtraction is an error in task programming or planning. The analogy is gross, but it exemplifies the necessity of arduous and exact planning of learning tasks.

The third type of child encountered has experienced many failures due to errors in expectations for task performances, and consequently the absence of benefits which are associated with successful school performance. Hence, this child often refuses to interact with a presented task, or when he does, completes it hurriedly and erroneously. In the classroom, the teacher may observe extreme forms of avoidance behaviors which range from low environmental interaction frequency, to high frequency of assaultive behaviors designed to escape from school related expectations for appropriate academic and social behaviors. For a child of this type, curriculum or task revisions and the addition of synthetic incentive conditions may have to be implemented to ensure task, child interaction. Only with many successes associated with this arrangement will the child be able to progress toward larger task presentations, and the phasing out of synthetic conditions. Natural benefits of successful learning are paired with the synthetic conditions until it is possible to phase out and eliminate the synthetic aspects of the total program.

Effective teaching is concerned with the three types of children, how to assist them in deleting behavior excesses and deficits, and procedures for preventing their occurrence. Other sections in this chapter discuss and describe these procedures.

Analysis of Behavior Excesses and Deficits

Excesses and deficits in children's behavior can be observed and described. Once this aspect is completed, an analysis of child, teacher, environment interaction must be instituted to determine the variables or environmental events which are maintaining behavior excesses and deficits.

Identifying behavior excess or deficit provides a description. It may be noted that a child is out of seat during an arithmetic lesson 70% of the time, or 7 times in a 30-minute period. At this juncture a functional analysis (Skinner, 1953) of behavior is necessary. What are the environmental events or variables that occur immediately before and after the identified behavior? These events function to determine the occurrence of the behavior at a specific time, place, and set of conditions. Teachers are instructed to be cognizant of the connective relationship between behavior and environmental events (Skinner, 1963). The very important concept of structure and its components is based upon the clarification of the relationship among events which occur before and after a described behavior (Haring and Phillips, 1962). Analyzing such connective relationships has culminated in the discovery of behavior principles which account for a considerable portion of human behavior, and with discerning use of them it is possible to provide a basis for predictability and lawfulness of behavior (Nurnberger, Ferster, and Brady, 1963). It is not asserted that behavior principles account for all behavior, but they do provide guidelines and procedures which can be used effectively to alter behaviors that interfere with adjustment and learning. Such principles can also be used to assist individuals in acquiring those behaviors that are necessary to maintain adequate adjustment and learning.

Many behaviors are learned, changed, and maintained by environmental events which occur before and after the behaviors. These environmental events may function to decrease or increase behavior. Effective teaching is concerned with procedures based upon behavior principles that can be instituted to change behavior, with the relationship between environmental events which occur before and after behavior, and the effect of these events upon behavior changes. A major emphasis is placed upon behavior which can be measured or counted, and is observable. This aspect does not deny the importance, contributions, and effects that central nervous system, chemical, and physiological interactions exert on behavior. At the present time, however, these variables are difficult to observe in a direct manner, and are not as available for alteration to produce behavior changes as are environmental events, considered in this context, that can be manipulated to change behavior excesses and deficits.

Changing or altering behavior requires systematic arrangements between behavior and environmental events that occur before and after behavior. The major objectives are to decrease behavior excesses, and replace behavior deficits with adaptive, adjustive behaviors. Concentration upon just one of the objectives when both must be considered represents only partial completion of the total behavior change process.

<u>Relevant aspects of environmental events</u>. Environmental events occur before and after a specific behavior. One example of an event occurring

before a behavior is a teacher's verbal direction to a child to pick up a book. The specific behavior of concern is the child picking up the book. Teacher comment of "thank you," or "that's fine" is an example of an event occurring after the specific behavior. A visual sequence is as follows:

<u>Before Event</u>	<u>Specific Behavior</u>	<u>After Event</u>
"Pick up the book"	Child picks up book	"Thank you"

<u>Before environmental events: Analysis</u>. These events may be referred to as before or antecedent (Lindsley, 1964) to a specific circumscribed behavior. If the event can be demonstrated to decrease the frequency or number of times a behavior occurs it may be labeled a decreasing stimulus. If, in relation to the visual sequence example, the child fails to pick up the book 8 times out of 9 requests, in comparison to picking up the book 9 times out of 9 requests when the request was preceded by the word "please," then the before event of "pick up the book" is functionally a decreasing stimulus. A further illustration of a decreasing stimulus is the presentation of a difficult reading task which reduces the probability of a child completing it because it is too difficult.

The before event may also function to increase behavior. If this can be demonstrated through behavior analysis, the event may be described as an increasing stimulus. The analogy cited for the demonstration of a decreasing stimulus may be reversed for determining a functional increasing stimulus. Teachers often observe one child making a fact at another child who then yells upon presentation of this before event. If the child yells consistently, the face making is an increasing stimulus. If the face making does not elicit a behavior from another child it still may be a before event but is not functionally related to another child's specific behavior.

There are many examples of before events which occur in the classroom. Requests from teachers, peers, custodians, and tasks which require verbal or written responses are just a few situations which can be cited. Concern with before events is directly related and relevant to task and curriculum planning which teachers must complete in developing behavior change programs for children with excesses and deficits. Before events may be used as increasing stimuli to ensure correct arithmetic problem completion, or as decreasing stimuli to reduce the frequency of yelling.

After environmental events: Analysis. The events may be described as after or subsequent (Lindsley, 1964) to a behavior. As is true with before events, after events may have functional effects upon a specific behavior. It can be demonstrated that after events will function to decrease the frequency of a behavior and may therefore be defined as a decreasing consequence. If "thank you" reduces the frequency or number of times the child will pick up the book upon request, then "thank you" functions as a decreasing consequence. A teacher may observe that the frequency of inappropriate talk outs in class decreases when each talk out results in a one-minute loss of recess; the one-minute loss is a decreasing consequence.

An after event may also function as an increasing consequence. Again the analogy of a child picking up a book is applicable. The after event "thank you" may be directly related to increasing the frequency of picking up the book when requested and therefore functions as an increasing consequence. Similarly, the addition of one minute to recess for each correct spelling word written may operate as an increasing consequence.

Before and after environmental events: Synthesis. It is quite possible and feasible to determine the functional relationships among before events, behavior, and after events (Lindsley, 1964). The teacher can keep the before event and the specific behavior constant and then vary the after event. If changing the after event is associated with a frequency change in the behavior specified, the after event may be described as a consequence. Varying only the before event, and observing any behavior changes may operate to define a functional stimulus. If the behavior frequency does not change, functional relationships are not demonstrated; the before and after variables are events and may not be defined as stimuli or consequences.

The implications for planning tasks for children with behavior excesses and deficits are apparent. Teachers can recognize the importance of stimulus, individual child behavior, and consequence as they plan child, environment transactions. Recognition that the variables described relate to aspects of curriculum planning, understanding of behavior principles, specifying behaviors to be changed, and implementing a total behavior change program based on these aspects delineate some of the attributes that an effective teacher of children with behavior disorders must possess.

Increasing vs. decreasing. Recognition that increasing a behavior may not always be good or positive, and that decreasing a behavior is not always bad or negative is necessary for effective planning of behavior change programs. Before either may be defined as negative or positive, the functional effects upon behavior must be ascertained. For example purposes consequences will be selected, represented visually as follows:

Consequences

Increasing Consequence	Decreasing Consequence
1. positive effect: word recognition is more accurate when time to build a model is contingent upon increases in word recognition accuracy.	1. positive effect: talk out frequency decreases when teacher and peers ignore it.
2. negative effect: talk out frequency increases when teacher and peers attend to it.	2. negative effect: rapid and completely correct word recognition task is followed by additional word recognition tasks.

If environmental events function to increase behavioral excesses which should be decreased, or further operate to maintain deficits which should be reduced or replaced, the effects are essentially negative. Reversing the order of events by decreasing excesses and building more appropriate behavior to replace behavior deficits represents a positive effect.

Negative and positive intervention programs to decrease behavior. A teacher may decide to decrease a behavior excess, and ensconced within that decision is a method to proceed in a negative or positive manner. If at all possible, the positive approach to intervention or behavior change should be implemented.

The teacher observes that a child is out of seat too frequently (excess) to be successful in task completion. Tasks do not refer just to academics, but may and should include social interaction too. The objective of the intervention program is to decrease inappropriate out of seat behavior, or reduce it to zero.

A negative approach to the problem involves applying a decreasing consequence to out of seat behavior. Each time the child leaves his chair without permission may result in a consequence such as loss of recess time, reduction of free time to read a comic book, or removal of the opportunity to work for increasing consequences through placement in time-out (Whelan, 1966a). The negative approach may be effective in that frequency of out of seat behavior decreases under the consequences specified.

A teacher may also approach the same problem in a positive manner. That is, the child becomes the recipient of increasing consequences for each segment of time that he spends at his desk engaged in task completion. He may earn points for each correct arithmetic problem, and earned points may be traded for time to engage in high probability behavior (Homme, 1966; Whelan, 1966a) such as listening to records.

Staying in the seat is an incompatible behavior with out of seat. The positive approach emphasizes the pleasurable increasing consequences accumulated for being in seat. The negative approach uses somewhat painful decreasing consequences for being out of seat without permission. The behavior observations and the graphic visual summary of the behavior change process will indicate a decrease in out of seat behavior with either method or approach. The behavioral effect is the same, but one was achieved positively, the other negatively.

The positive approach has the added advantage of requiring scrutiny of task expectations presented to a child. If a child is leaving his chair and desk frequently, it may be evoked by inappropriate program planning. Therefore, a positive approach emphasizes (1) an analysis of learning tasks, (2) possible revision of learning tasks, and (3) providing consequences for completion of the tasks.

Behavior Analysis and Procedures for Changing Behavior

The effective teacher of children with behavior disorders (excesses and deficits) needs to understand, recognize, and implement behavior principles. This understanding must not only be reflected in verbal statements relevant to behavior principles and behavior analysis, but it must be translated into procedures which are effective with children. Verbalization of behavior principles does not necessarily guarantee correct, systematic and sensitive subsequent application. A teacher must be cognizant that the behavior principles and processes described are relatively simple to understand and are few in number, but applying them systematically, accurately and effectively can be a very complex endeavor.

The responsibility assumed by educators for children in a school situation requires a commitment and investment to ensure that learning experiences are planned and implemented in the most positive manner possible.

If a teacher is concerned that a child is not progressing adequately, or that some aspect of current behavior is interfering with classroom adjustment, procedures for changing behavior should be initiated. Again, the procedural processes are simply described and understood, but must be implemented with exactness and careful attention to organizational planning.

Five procedural steps are involved in planning a behavior change program. Each step in the process is vital to correct implementation of a behavior change program. The five steps are as follows:

1. Behavior selection
2. Behavior recording
3. Behavior events
4. Behavior intervention variables
5. Behavior change evaluation

Steps 1 and 3 are associated with a segment in a behavior change program which is labeled "before intervention." Step 4 is the point in the program during which an environmental event or variable is introduced to foster behavior change in the desired direction. This segment of the program is defined as "during intervention." Step 5 is relevant to assessing the behavior effects observed during the period when the intervention variable is in effect in comparison to the status of the behavior before intervention. Evaluation also continues beyond the "during intervention" segment to a segment described as "after intervention." That is, it is important to discern the impact which removal of an intervention variable will exert on the behavior. Will the behavior continue at the same level recorded during the intervention segment, or will it decrease or increase? Step 2, behavior recording, must be operational and continued throughout the entire behavior change program. The five steps necessary to initiate and implement behavior change programs are divided into three important segments defined as (1) before intervention, (2) during intervention, and (3) after intervention. Steps 1 and 3 are identified with segment 1; steps 4 with segment 2; and steps 2 and 5 are associated with all segments of a behavior change program.

Behavior selection. The first step in instituting a behavior change program is to pick or choose a behavior or behaviors which need to be changed. It may be a behavior that needs to be increased (correct word recognition) or decreased (out of seat). The behavior must be defined and described in a circumscribed manner. That is, it must be described precisely and specifically. For example, a teacher when asked to describe a concern pertinent to a particular child may mention behavior episodes by using words such as hyperactive and aggressive. The problem is to assist the teacher in placing such words in a context that will provide more exact behavior descriptions. Hyperactive may be translated into out of seat, turning around when attention to work is suspected, and running to teacher's desk. Aggressive behavior may mean hitting other children in the classroom, halls, and on the playground. Most teachers can provide exact descriptions of behaviors which are interfering with child, teacher, environment interactions.

Behavior recording. Once the behavior has been selected, and it is translated into a behavior unit or cycle which can be observed, it should be counted and recorded. If a teacher is concerned that a child is talking out frequently, interfering with task completion and disrupting other children, the number of talk outs should be counted. After several days of recording, the teacher can state very specifically that talk outs during arithmetic period range from 3 to 9 times. If the teacher, with the evidence available, still believes that talk outs are too numerous, behavior change procedures can be planned.

The initial recording of the specified behavior should be accomplished in a hidden manner. That is, the child should not be aware that he is being observed or that some aspect of his behavior is being counted. Child awareness that a behavior is being observed and counted may result in that behavior being affected. For example, if a child learns that the teacher is counting inappropriate talk outs, that behavior may decrease in frequency or be reduced to zero. Awareness of behavior counting and recording may be a sufficient variable to change the specific behavior in the desired or undesired direction. The crucial aspect at this juncture in behavior change procedures is that the teacher must be aware of the possible effects child awareness of behavior recording may exert. If the teacher is not aware, a subsequent behavior intervention variable that is introduced may be erroneously identified as the one that produced the behavior change. In essence, child awareness of behavior observation and recording may function as the intervention variable, and further investigation in regard to the specified behavior may not be warranted.

The period of observing and counting a selected behavior has been referred to as a baseline period (Ullmann and Krasner, 1965). It may also be identified as the before intervention period. The reason for counting or quantifying behavior is to provide a reference line that can be used to compare intervention procedures when they are implemented after the before intervention period. Subsequent to the before period, behavior change procedures are initiated. At this point the teacher may simply inform the child that his behavior is being counted. Further observations will enable the teacher to determine if this intervention variable is effective in providing behavior change in the desired direction. The recording procedure may be sufficient to produce behavior change and it is important for teachers to be aware of and recognize the implications of this effect or function.

Behavior observation methods. When teachers are requested to observe and count behavior, one frequent and crucial question asked pertains to the time element. Even in a special class with 8 children, the teacher may not have the time to devote to individual child behavior counting. While observing one child, others may be engaging in behaviors far more deleterious than the one observed.

However, observation methods do exist which can be realistically applied by teachers. If the specified behavior is focused upon speed and accuracy in arithmetic task completion, the teacher need only write down the time the child started the task and the time when it was completed. Examination of the completed paper reveals the per cent correct, number and type of errors, and the time consumed to complete the task. This system may be used with most materials that require verbal or oral responses, but the teacher must be available to record correct and error responses while listening to the child's performance.

For social behaviors such as out of seat, talk outs, hitting, etc., the teacher may use several systems. The least time consuming one is a periodic or aperiodic glance at the child and observing if he is exhibiting the specified behavior or not. For example, the teacher may decide that for a one-hour period 10 behavior spot checks will be completed. These checks may occur every 6 minutes or on the average of every 6 minutes. It may be observed that the specified behavior was occurring during 8 of the spot checks. With additional personnel available for behavior observations, such checks may be made every 20 seconds or once per minute. However, the calculation of behavior occurrences remains the same.

While the spot check system is not too time consuming, it does not provide as much information about the selected behavior as the time sample system. The teacher may decide that the child will be observed continuously for 6 15-minute periods spaced throughout the school day. During each of these periods, the teacher observes and records all occurrences of the specified behavior. This system provides a numerical count of the behavior during 90 minutes of daily observational time.

The third system is a continuation of the time sample approach. It is continuous recording of the specified behavior for the entire school day. Realistically, teachers may not be able to conduct this type of observation and recording system because of the time which must be devoted. It does provide more information relevant to the selected behavior, but it may require an additional person to observe and record the behavior. Additional personnel, however, may not be available to most teachers.

Behavior recording and counting methods. Once the teacher has devised a behavior observation and counting system, a method of recording the behavior count must be instituted. One method uses frequency, or number of behavior counts per unit of time. If the child talks out 30 times in a 15-minute period, the frequency is two per minute, or 120 per hour.

A per cent method may be used entirely or in conjunction with frequency. For an arithmetic task, the number of correct problems completed per minute and the per cent accuracy must be calculated. The child, depending upon the complexity of the task, may be completing two correct problems per minute and have an accuracy score of 70%. Another child may only be completing one problem per minute with 95% accuracy. These two sources of information can assist the teacher in decisions regarding establishing a level of performance. If 90% is the criterion, perhaps some reduction in frequency will be necessary. Or the teacher may decide to use the most efficient combination of frequency and per cent, e.g., 90% accuracy and a frequency of one problem per minute. If 90% accuracy in task completion is the single criterion, frequency can be allowed to vary as a function of task complexity. Time elements may be based upon the length and complexity of tasks, but the per cent accuracy score remains fixed. If per cent accuracy decreases below 90%, the teacher may use this information to revise task expectations as they relate to before events or stimulus variables.

Numerical counting is another method which the teacher can use effectively in recording a selected behavior. This can be instituted in situations where the teacher desires to obtain a total incidence figure for a one-hour period or for a six-hour school day. For example, during a 60-minute lunch hour the child may be observed hitting peers five times. Numerical counting may be used alone, but it is also used in determining frequency and per cent measures.

Some behaviors, in terms of the time element and aspects inherent in the specified behavior, may be more effectively recorded or counted by the duration method. For example, a teacher may select rocking as a behavior which should be decreased because it interferes with a reading task. It is extremely difficult to count the number of times the child's back hits the chair unless another observer is present or an automatic counting device can be attached to the chair back. This system would provide the actual number of rocking cycles. However, if the teacher recorded the starting and stopping times of a rocking episode, this would provide a duration measure in number of minutes the rocking behavior occurred. The child may have two 5-minute episodes of rocking in a 20-minute period designated for a reading task. Ten minutes out of 20 minutes were spent rocking in the chair. This duration count could then be expressed in a minute system per each 20-minute segment of time or it could be indicated with the per cent system. For each 20-minute segment, the child spends 50% of the time engaged in rocking.

As indicated, behavior may be recorded and counted by several methods. In selecting the method to be used, it is necessary to relate that decision to the specified behavior that is to be changed. Finger sucking, rocking, etc. may be recorded more efficiently by the duration method. Hitting other children, out of seat, academic performance, etc. can be recorded numerically and then translated into a frequency or per cent measure. The crucial aspect is that the counting method should be selected on the basis of the specified behavior, instead of forcing the behavior into a predetermined counting method. Counting and recording methods should function to enhance a realistic appraisal and assessment of the selected behavior. It is technically possible to count the number of finger sucks and convert those to a frequency measure, but it is probably not feasible or desirable for a teacher to devote that much time to the counting method. If the teacher desires to decrease finger sucking, a duration measure can indicate the success or failure of a behavior change program.

Behavior counting in the classroom can be achieved by using several sources. The teacher is usually involved in the recording of a behavior either directly or in supervising another individual who is recording. A child may record his own behavior by writing the start and stop time on the academic tasks and, with the assistance of the teacher, writing down the number of correct responses. Child or self recording may function as a behavior change variable in that once the child discerns the behavior to be changed and is actively engaged in a self behavior change procedure, the behavior may change in the desired direction. However, if the child counts and records self behaviors the teacher must assess the reliability and accuracy of the recordings. Peers can be used to record the behavior of other children but again the importance of accuracy must be stressed. Some automatic recording devices can be used. At the Children's Rehabilitation Unit one teacher uses a time clock. Before beginning a task, the paper is placed in the time clock and stamped. The same process is used when the task is completed. This system frees the teacher from writing down the times for the children, plus provides an accurate time check on task performance.

Teachers, self, peers and automatic devices can be used to provide a record or count of specified behaviors. The counts can be recorded several ways. Some teachers use a golf counter to record a behavior when it occurs. Others use a small pad of paper to record the count. The specified behavior is written at the top of the pad, and each behavior occurrence is tallied (////). Teacher-devised ditto forms, especially for academic tasks, can be used by teachers and children. For example, the form may contain segments or blocks which indicate number of tasks expected to be completed, time started and time stopped, number correct, number incorrect, and per cent correct.

Behavior events. Concurrent with the behavior recording the teacher should attempt to identify the associated conditions which affect the specified behavior. That is, what are the before and after events which possibly function to maintain the behavior at an undesirable level of occurrence. If one child talks out frequently, and the peer group attends to that, this after event may be functioning as an increasing consequence that keeps the behavior at a high frequency. Similarly, if a child's performance in reading is decreasing, it may be due to the before event. The before event or material may be presented in progressions that are too large for the child to comprehend. Identifying the before and after events can assist in deciding what aspects of the child, teacher, environment interaction need to be changed.

Behavior intervention variables. The next step in the behavior change program is to determine or select a variable that will be introduced to change a specified behavior in the desired direction. One of the associated events may be eliminated or changed. A new variable such as a change in the after event may be introduced as a portion of the during intervention segment.

As an example of this process, two teachers at the Children's Rehabilitation Unit (Mrs. Ruth Mulder and Miss Ann Jarvis) decided that a child was making too many errors and not completing enough of the arithmetic assignment. Instead of instituting a behavior change program based on points which could be traded for free time, or more extrinsic events such as candy, the teachers told the child that each arithmetic problem completed correctly would be worth time which could be spent in the arithmetic laboratory. This laboratory consisted of an adding machine and some problems which could be worked using the machine. The behavior recording indicated an increase in accuracy and in total performance completion. An after event was determined to function as an increasing consequence. The child's performance increased as a function of the contingent use of the arithmetic laboratory. In essence, the child met expectations for performance in order to gain access to the increasing consequence. Classrooms are virtually banks of potential intervention variables. The opportunities for creative and systematic use of before and after events defy description or cataloging.

Two other staff members at the Children's Rehabilitation Unit (Dr. Roger Kroth and Dr. Gene Plank) used children in classes as increasing consequences for other children. An adolescent boy met expectations for classroom performance for the opportunity to tutor a younger boy. Before the younger boy could gain access to the tutor, previously assigned work had to be completed adequately. Both boys profited in that the tutorial interaction was task completion oriented, and that completion of other relevant tasks was related to the segment of the behavior change program.

Behavior change evaluation. The next phase in the procedures for changing behavior is to evaluate the effects of the variable introduced by continuing to observe and record the specified behavior. If the behavior changes in the desired direction, the behavior change program may be evaluated as successful. However, if the selected behavior does not change or changes in the opposite direction than what is desired or predicted, the program is not successful. A teacher may wish to increase spelling accuracy by instituting an intervention variable of one stick of gum for each word spelled correctly. If, in subsequent lessons, the accuracy is not improved, the gum operates as only an after event. It does not function to change the specified behavior. However, the gum could operate to decrease spelling performance, and therefore functions as a decreasing consequence.

Continuation of the behavior observation and recording enables the teacher to evaluate the effectiveness of the behavior change program.

This evaluation further provides direct evidence which can assist in decision making regarding termination of the intervention segment, or initiation of a new intervention variable if the first one was not effective.

These five procedures involved in planning behavior change programs should be followed by a teacher who decides to assist children in changing behavior excesses and deficits. In a summary form the procedures are as follows:

1. Pick or select the behavior to be changed.
2. Record the occurrence of the behavior by observation and counting.
3. Identify the associated behavior events (before and after) which may be affecting the current level or frequency of the selected behavior.
4. Determine and introduce an intervention variable which may function to influence the behavior in the desired direction.
5. Evaluate the effectiveness of the variable introduced as a segment of a behavior change program in order to determine if the behavior is changing in the anticipated or selected direction.

Behavior Display

Following the five procedures discussed and described for changing behavior will provide a substantial amount of information pertaining to a child's behavior that has been specified and selected for change. The teacher will have accumulated, minute, hourly, and daily records of the behavior over a range of perhaps five to twenty school days. Actual number of school days, tutorial sessions, etc. during which the specified behavior is observed, counted, and recorded is dependent upon the type of behavior to be changed.

The behavior measures, e.g., frequency, counts, per cent, number correct and incorrect, etc., are available in the records which the teacher has compiled. Concurrent with the behavior observation and recording, the teacher should keep a summary visual chart of the behavior. The type of visual display will depend on the behavior selected, and the method or system chosen to observe, count, and record the behavior.

It is important to recognize that a visual display or graph provides a summary of the total process that involves implementation of the five procedures used to foster behavior change. It provides a summary record of a behavior change program and, even more pertinent, it is a record of the effects of a behavior change program.

The visual display may be placed upon ordinary graph paper which is readily available in classroom situations. It should be divided into segments which correspond with behavior change steps or procedures. An example of this is portrayed in Figure I.

The teacher, (Mrs. Nina White, Wichita, Kansas) decided that the boy was exhibiting too many errors in oral reading. As indicated in the first segment of the graph (Before Intervention) the per cent correct was quite low. At this juncture of the behavior change process the teacher had to make a decision regarding the exact form which the intervention segment would entail. For example, a decision had to be made regarding a change in before or after events. Upon examination of the graph, one possible speculation is that the material read is too difficult and therefore expectations for performance were unrealistic. However, the teacher believed that the material was appropriate and the boy was avoiding interaction with task expectations. Consequently, the decision was made to introduce the variable of opportunity to earn car parts contingent upon a specified performance criterion.

As indicated on the second segment of the graph (During Intervention) the after event was demonstrated to function as an increasing consequence. During the first three days of the intervention segment, the boy did not achieve the performance level which was necessary to obtain a car part. This aspect, which many children with behavior disorders exhibit, has been described as a shaping or reality testing stage (Haring and Whelan, 1965). That is, the boy was testing the teacher to determine if there would be a match between verbalization of intervention conditions, and actual implementation of the conditions.

Segment three (After Intervention) indicates that the boy continued to read orally at an acceptable performance level. This segment of the total behavior change program is vital in that it allows the teacher to

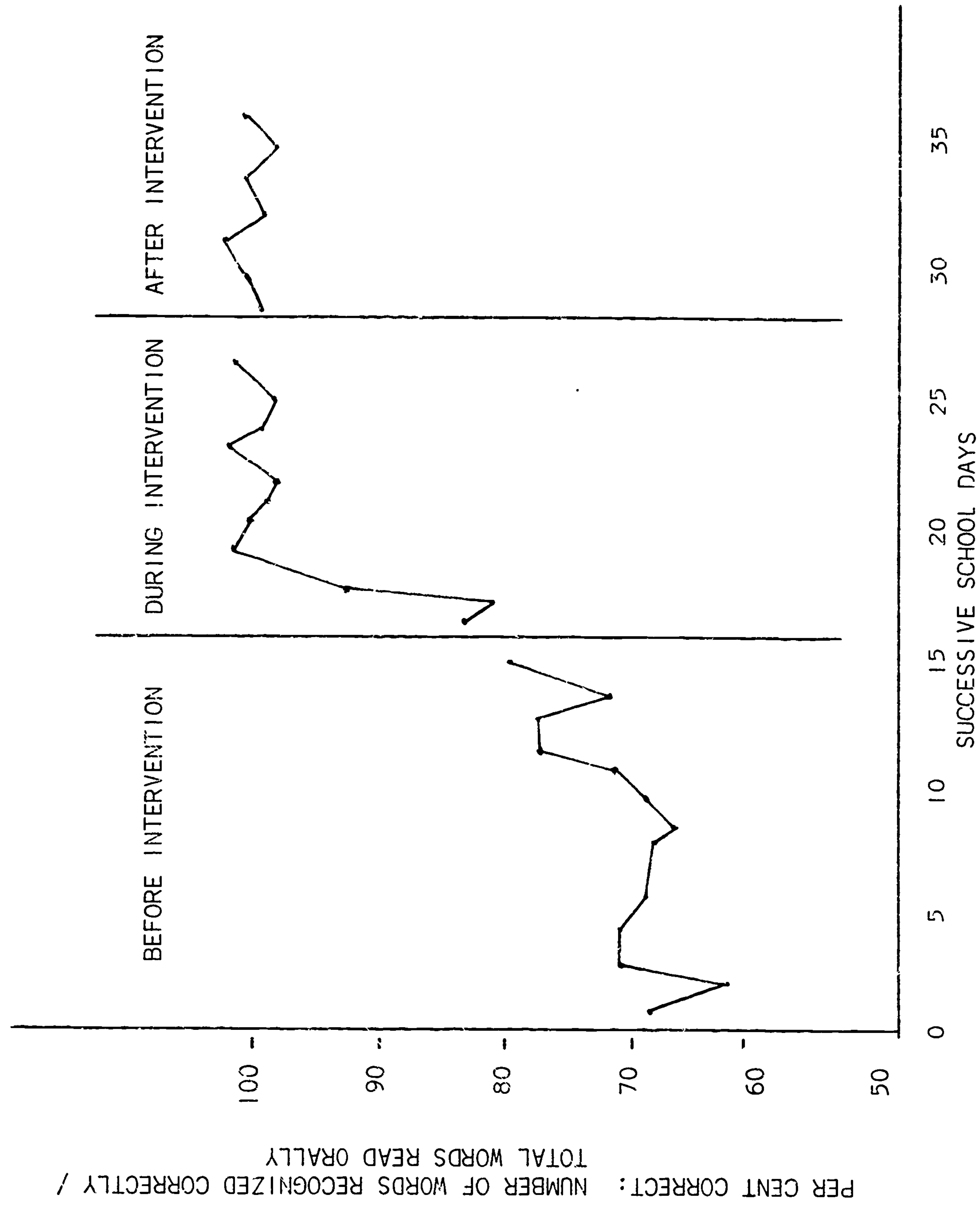


Figure 1: The effects of an intervention program designed to increase correct word recognition in oral reading. Subject was a 13-year-old emotionally disturbed male. Intervention program was the opportunity to obtain one car part of a plastic model for a per cent word recognition score of 95 or higher.

determine if appropriate task performance is maintained after an intervention variable is removed. If the per cent correct had decreased, the intervention program would probably have to be reinstituted since it was terminated too soon. Evidence that per cent correct decreased to before intervention level may indicate that the after events normally available in the classroom, such as praise, stars, self gratification from achievement, etc., are not functioning to maintain the specified behavior of the desired level. However, as displayed in Figure 1, the performance was maintained. At this stage the teacher may only have to record the child's performance on an aperiodic basis of twice a week to ascertain if the behavior is being maintained at an appropriate level.

Behavior Acquisition and Maintenance

The importance of differentiating the difference between acquisition and maintenance has been stressed (Lindsley, 1964; Whelan and Haring, 1966a; Whelan, 1966b). In the case displayed in Figure 1 behavior acquisition occurred during the intervention segment when earning car parts was contingent upon an established performance criterion. Behavior records during the after intervention segment revealed that the desired behavior was being maintained by other environmental events which were functioning as consequences.

Appropriate behavior should be maintained. If teachers do not attend to this important aspect, acceptable levels of behavior may decrease. The environmental events or conditions necessary to acquire behavior are different from those necessary for behavior maintenance. The main difference, for example, resides in the scheduling of consequences. During behavior acquisition, each arithmetic problem completed correctly may earn one car part. Over time that schedule of performance to consequences may be changed from 1:1 to a 5:1 ratio. That is, the child must complete five problems correctly for each car part.

After the intervention segment is completed and, for this example, the car parts have been earned, the teacher should continue to record the specified behavior to determine if other environmental events are maintaining it at an adequate performance level. During behavior maintenance the teacher may notice changes in consequences necessary to sustain behavior at an appropriate level. Extrinsic or synthetic consequences such as car parts, toys, candy, etc. may have to be used to enable a child to acquire a behavior. Concurrent with this aspect, intrinsic or natural environmental events such as teacher approval, self satisfaction with achievement, etc. which did not function as consequences in the past are paired and obtained with the extrinsic consequences. If, after the acquisition segment of the intervention program is terminated and the extrinsic consequences are gradually diminished, the specified behavior remains at criterion level the behavior is being maintained by environmental events which previously were not functioning in this manner.

Behavior Change Questions

Educators of children with behavior excesses and deficits must know and understand the principles and procedures upon which behavior change programs are based. In addition, daily systematic and correct application of the principles and procedures is necessary if children with behavior disorders are to obtain assistance in managing and organizing their rather disorganized, painful, debilitating existence.

In essence, the effective teacher of children with behavior excesses and deficits must post four questions relevant to behavior change programs implemented. The questions are:

1. What are the terminal behaviors or competencies that the child is expected to acquire and exhibit?
2. What are the learning tasks, experiences and environmental events (before and after) that must be planned in order to ensure sequential and continuous progress toward acquisition of terminal competencies?
3. What are the behavior observation, counting, and recording procedures which are instituted to evaluate progress toward the acquisition of terminal competencies?
4. Will the child continue to maintain an acceptable or appropriate level of terminal competencies learned during the behavior change program in situations different from the conditions in which they were acquired?

The teacher must be cognizant of these four basic questions during the implementation of a behavior change program. Behavior analysis and the procedures used to foster changes in children's behavior are based upon the necessity of obtaining specific and accurate answers to the four basic questions. Previous sections discussed and described the methodology and rationale of applying procedures which are designed to provide adequate answers to the questions.

Implications of Effective Teaching

A plethora of implications may be listed and discussed relevant to the systematic utilization of behavior change procedures. Using these procedures requires a realistic degree of self evaluation and criticism by educators. If a program does not function to produce positive behavior change it must be discarded and replaced with another program which in turn must be evaluated. In these instances, it is apparent that failure to achieve objectives is not blamed on the child. In reality, all that can be stated is that the child has not acquired competencies under the conditions provided to obtain such competencies; additional effort is necessary.

As previously discussed, it is important to recognize that the behavior change procedures described are simple in principle, but complex in requiring systematic, accurate, sensitive and effective application. These procedures provide educators with a valuable and powerful methodology which can be implemented to assist children with behavior excesses and deficits. However, the method is neutral; it is neither good or bad. It is selected by individuals who may use the procedures accurately or inaccurately, and who may apply them for motive that can be labeled negative or positive.

Behavior change procedures are not a panacea. They do, however, provide teachers and children with a methodology that, if applied correctly, will result in beneficial progress toward realistic school performance, adaptive behavior, self control, and adjustment. Criticism may be expressed in terms of statements that the procedures are not new and that educators have used them for years. There is a large element of veracity in such statements. Most teachers use motivational incentives such as stars, progress charts, grades, social approval, verbal encouragement, etc., but the important aspect that is often neglected is the exactitude and systematic utilization of task expectatio.s, and environmental consequences which is vital if behavic- excesses and deficits are to be changed in an effective manner. Behavior change procedures, if followed precisely, do provide for effective application. It is also true that behavior change procedures are not an entirely recent innovation as partially illustrated in the following passage:

> To leave a child wholly to his own inclinations in reading is as absurd as to send him to take honey from a swarm of angry bees and not expect him to be stung. To supply him with honey, all that he wants, at all times and without exertion to himself, is to clog his taste and destroy his appetite. We must see that he is led to look for the sweet, taught to recognize it when he finds it, and to extract it from the comb. He will enjoy working to get it. On the other hand, he must not be sent where the reward is too difficult to find and secure, lest he become discouraged and cease to work (Sylvester, 1909).

The passage describes most elements involved in applying behavior change procedures. The major tasks are to describe those elements exactly, apply them in a systematic manner to children with behavior disorders, and translate behavior principles known for years into a functional methodology for teachers and children. Ordered learning situations, procedures utilized to foster appropriate degrees of self esteem and satisfaction, and positive evaluations which occur from successful task completion, function to alleviate the myriad excesses and deficits that are observed in children with behavior disorders.

References

Bower, E. M. Early identification of emotionally handicapped children in school. Springfield, Ill.: Charles C. Thomas, 1960.

Eysenck, H. J. The effects of psychotherapy. In H. J. Eysenck (Ed.), Handbook of abnormal psychology. New York: Basic Books, 1961.

Gallagher, P. A. A group contingency study: Observation and recording of students' behavior. Paper presented at the meeting of the International Council for Exceptional Children, New York, April, 1968.

Haring, N. G. and Phillips, E. L. Educating emotionally disturbed children. New York: McGraw-Hill, 1962.

Haring, N. G. and Whelan, R.J. Experimental methods in education and management. In N. Long, W. Morse, and R. Newman (Eds.) Conflict in the classroom. Belmont, Calif.: Wadsworth, 1965.

Homme, L. E. Human motivation and the environment. In N. G. Haring and R. J. Whelan (Eds.), The learning environment: Relationship to behavior modification and implications for special education. Lawrence, Kans.: University of Kansas Publication, 1966.

Kanner, L. Autistic disturbances of affective contact. Nervous Child, 1943, 2, 217-250.

Lindsley, O. R. Direct measurement and prosthesis of retarded behavior. Journal of Education, 1964, 147, 62-81.

Nurnberger, J., Ferster, C. and Brady, J. An introduction to the science of human behavior. New York: Appleton-Century-Crofts, 1963.

Overshiner, B. Problems of school children as seen by themselves. Unpublished master's project, Children's Rehabilitation Unit, University of Kansas Medical Center, 1968.

Pate, J. E. Emotionally disturbed and socially maladjusted children. In L. M. Dunn (Ed.), Exceptional children in the schools. New York: Holt, Rinehart, and Winston, 1963.

Phillips, E. L., Wiener, D. N. and Haring, N. G. Discipline, achievement, and mental health. Englewood Cliffs, N. J.: Prentice-Hall, 1960.

Rimland, B. Infantile autism. New York: Appleton-Century-Crofts, 1964.

Skinner, B. F. Science and human behavior. New York: Macmillan, 1953.

Skinner, B. F. Operant behavior. American Psychologist, 1963, 18, 503-515.

Skinner, B. F. The technology of teaching. New York: Appleton-Century-Crofts, 1968.

Sylvester, C. H. Guide to journeys through bookland. Chicago: Bellows-Reeve, 1909.

Ullmann, L. P. and Krasner, L. (Eds.) Case studies in behavior modification. New York: Holt, Rinehart, and Winston, 1965.

Whelan, R. J. and Haring, N. G. Modification and maintenance of behavior through systematic application of consequences. Exceptional Children, 1966, 32, 281-289. (a)

Whelan, R. J. The relevance of behavior modification procedures for teachers of emotionally disturbed children. In P. Knoblock (Ed.), Intervention approaches in educating disturbed children. Syracuse, N.Y.: Syracuse University Press, 1966. (b)

Whelan, R. J., Gallagher, P. A., Grosenick, J. K. and Kroth, R. L. Enhancing children's learning. Unpublished manuscript, Children's Rehabilitation Unit, University of Kansas Medical Center, 1968.

CHAPTER IX DEVELOPING COOPERATIVE SOCIAL BEHAVIOR

L. J. Peter

Introduction

Prescriptive Teaching (Peter, 1965) offers an effective approach to the implementation of educational programs to cooperative social behavior in disturbed children. This methodology is being employed in the Prescriptive Teaching Center at the University of Southern California with autistic, severely emotionally disturbed, neurologically impaired, and educationally handicapped children.

An advantage of prescriptive teaching to learning and behavior disorders is that it functions within the accepted structure of the educational process. Although it requires considerably more specificity and precision than the methods generally employed in education, the method is consistent with the professional role of the teacher and with educational theories and methodologies.

Prescriptive teaching begins by defining the child's present status. General educational objectives are translated into specific terminal goals and these become the behavioral objectives. The en route objectives, based on developmental psychology, provide a sound curriculum for social or academic learning. In addition, the methods and materials are systematically arranged to achieve the stated objectives.

The rationale of the approach used at the Center is presented here with a discussion of a number of cases involving problems in social behavior.

Rationale

Most children achieve the most complex learning, the ability to speak their native tongue, without formal instruction. This is because it is the nature of the human organism to make order and to attach meaning to experience. Much learning is the result of incidental or informal experiences rather than carefully structured or organized events. In the regular classroom, the child relies to a great extent on incidental learning.

The educationally handicapped child with developmental problems, learning disabilities, emotional disturbance, or social or behavioral disorders fails to benefit from the regular classroom situation because his ability to learn is impaired. Unlike the normal child, his incidental learning cannot be relied on. He becomes confused, disorganized, or develops self-defeating behaviors.

A therapeutic solution to this problem is to rely less on incidental learning and to structure the educational experiences for the child so that appropriate progressive changes in behavior are elicited and reinforced. This requires careful attention to the specifics in the teaching-learning process.

Prescriptive teaching is a means of determining these specifics and evaluating them in terms of educational relevance. Once they have been derived, they must be realized by an effective, therapeutic educational program.

Structure of the Teaching-Learning Process

Much of the information available relating to the child has educational relevance, but the degree of relevance can be determined by using the prescriptive teaching model to translate diagnostic information into educational terms (Peter, 1965). The educationally relevant variables are then organized and carried out through the educational process. To achieve this, the focus is directed to five fundamental aspects of the structure of the educational process: (1) the child's entering behavior, (2) the terminal objectives, (3) the developmental curriculum, (4) the eliciting events, and (5) the reinforcing events. The first three are part of the diagnostic teaching process while the last two are therapeutic procedures.

Diagnostic Teaching

The diagnostic teaching process determines the child's actual functioning level and establishes his responsiveness to a variety of stimulus and reinforcement events.

Entering behavior. A therapeutic educational program begins with the behavior the child presents on entering the program. The concept of "starting where the child is" has had wide acceptance in educational literature. This entering behavior is also termed initial behavior, present behavior, or baseline behavior.

Learning is defined as the progressive changes in behavior resulting from experience. In evaluating present behavior in terms of the learning process, it is necessary to see it in relation to the experiences producing the behavior. These experiences can be categorized as events occurring prior to behavior--eliciting or stimulus events (S), and those subsequent to the behavior--reinforcing or supporting events (R). For example, in evaluating a child's behavior (B), we must realize that the behavior occurs in an environment. A study of the child's attention span or study-related behavior reveals that his attention depends on his responsiveness to the task at hand (S) and to the support or reinforcement he derives from attenting to the task (B). This support may be in the form of external reinforcement (R) such as praise, smiles from the teacher, or food. The child's entering behavior, therefore, is evaluated as a behavior (B) that can be elicited (S) and reinforced (R). This can be shown symbolically: $S \rightarrow B \leftarrow R$.

The child's failure may be seen in any one of these three components or in a combination. The neurologically impaired child may be unable to produce appropriate reading behavior no matter how simple the stimuius events. For such a child, the teacher can provide more reading readiness activities or she can increase the stimulus value of the instructional materials. Another child may not respond to the reinforcements in a regular learning situation, but he may respond to food reinforcements or other more primitive material reinforcers. A hierarchy of reinforcements will be discussed later.

The child's entering behavior can be classified as motor, perceptual, language, social, cognitive, and problem solving. These provide a profile from which decisions of priority for remediation are made. The eliciting and reinforcing events on which contingencies must be evaluated in order to provide realistic meaning to the diagnosis is the content to determine the study of the child's entering behavior.

Terminal objectives. When the child's entering behavior is described by the above means, the terminal objectives for therapeutic teaching are usually obvious. For example, consider a child in grade four who does not respond appropriately to grade four reading material where other aspects of his development are consistent with fourth grade placement. The terminal objective for him would be to read grade four literature. This also defines the stimulus material as the regular grade four reading material. In other words, a specific objective for this child is for him to read in a regular classroom under normal circumstances with the kind of reinforcement schedule provided.

A child who exhibits behavior problems in the classroom and requires intense individual supervision at first would have as a terminal objective that he behave appropriately in a normal classroom situation with the normal stimuli which would include many irrelevant and distracting events. He must be able to attend to a task in this normal environment with the usual reinforcing contingencies. The terminal objectives can be described, therefore, with symbols similar to those of the initial evaluation. The terminal objectives are that the child will respond to the environmental stimuli and reinforcements with appropriate adaptive behavior. This is an increase in the child's behavior repertoire. The child is now able to read a long passage, solve more complex problems, maintain study-related behaviors, attend to a task, make finer discriminations, and behave in a socially acceptable manner. Behavioral objectives should be described in terms of observable or measurable behaviors.

Developmental curriculum. With a beginning and a goal defined for the therapeutic program, a question arises as to the developmental steps required to move from the entering behavior to the terminal objectives. Studies of child development, particularly in motor, language, cognitive, and social areas, provide guidelines. Many appropriate programs, such as curriculum guides, programmed instruction, and language development kits, are available. Experienced teachers are frequently effective in deciding on the steps and on the teaching methods and materials that will elicit these developmental steps in the child. For example, developmental reading programs provide the steps that might be appropriate for the fourth-grade child described.

Another case concerned a hyperactive five-year-old boy with severe behavior disorganization. He rolled around on the floor in the kindergarten room. He became excited by mild eliciting events. A light touch from another child resulted in violent behavior. His attention to a learning task was brief or fleeting. The boy was removed from the group and provided with a carefully structured sequence of learning tasks related to his behavior problem. Task one was to sit in his seat; two, to attend to a task; three, to complete a very short task; four, to complete a longer task; and five, to perform the task back in the classroom in an environment with normal distractions.

In order to begin teaching effectively, we need only know the first appropriate developmental step. But once this first step is achieved, we need to know the following step. It is not essential that all the steps be clearly defined before we attend to the en route objectives and the instructional sequence leading to their attainment.

Social Development

A knowledge of the relevant aspects of child development is valuable to determine en route objectives, but the remedial-educational model must provide for more specificity and flexibility, so that a task may be translated into small steps when a particular learning difficulty arises.

A general picture of social development follows:

<u>Birth to one year</u>. During the first month the infant may be soothed by being held or rocked, but this becomes more evident during the second or third month.

In the second and third months the development of visual fixation takes place. The child fixates on mother's face at close range and later smiles at a face.

Four four to six months he recognizes mother and distinguishes between familiar persons and strangers. He no longer smiles indiscriminately.

From seven to nine months he responds to simple social interactions and enjoys surprise games, such as hiding and peek-a-boo.

Between ten and twelve months he responds to his name, waves good-bye, plays pat-a-cake, gives and takes objects, and responds to "no."

One to two years. Between twelve and eighteen months he obeys simple commands, repeats a few words, feeds himself, is interested in his image in a mirror, and imitates many simple behaviors.

From eighteen to twenty-four months he expands the previous repertoire, the main new addition being that he may do the opposite of what he is told. He can actively cooperate in dressing himself. He holds a glass, usually with both hands, and drinks without much spilling. Play is mostly solitary. He likes to have other children around, but not for play; he may push them around as he does objects. He claims some objects as his own. He shows toys or offers them as a means of social contact.

Two to three years. His play is on a parallel level. He does the same things as other children but he does not play cooperatively with them. He uses "I," "me," and "you" in talking. He seeks adult commendation for correct behavior.

Three to four years. He initiates cooperative play, including sharing and taking turns. He uses "we" in talking. He imitates parents and models his behavior after the same-sex parent. He begins to ask questions.

Four to five years. He seeks out children for play. He becomes competitive.

Five to six years. He plans activities independent of his parents and plays authority roles. He prefers sex-appropriate activities.

Six to eleven years. He participates in more organized social activities including clubs, school, and games.

Twelve years and on. He begins dating and becomes involved in very complex social activities.

Although many details are lacking in the above, it can provide a basis for developing a sequence of social behavioral objectives.

Therapeutic Teaching

Learning occurs within the child. Teaching consists of providing events to occur before or after behavior of the learner. If these events are effective elicitors and reinforcers of progressive behavioral change, they constitute the teaching program.

Eliciting events. The events immediately preceding behavior are influential in eliciting developmental changes. They have traditionally been the major focus on teaching methods. The more appropriate the method or stimulus material is to the desired behavior, the greater the probability that it will elicit the progressive behavioral change. For example, in the case of the five-year-old described earlier, the first stimulus event was to guide the child to sit in his seat; the second was to provide him with a simple task, placing pegs in a pegboard; the third was to give him a complete set of pegs; the fourth, a longer task, required many more pegs in a more complex pegboard task; and the fifth was the same as the fourth task, given in a regular classroom setting.

Reading materials start with identifying common objects in the room, progressing to the use of these words in simple stories in a controlled vocabulary, and then moving on to the structural and phonetic analysis of the word, thus providing a developmental sequence of eliciting events.

When a child does not respond favorably to a stimulus, the response can be altered or remedied through reconditioning. A child with a negative emotional reaction to school can be reconditioned to react favorably if school and school attendance are associated with those things he enjoys.

Reinforcing events. Teachers have always used reinforcement, but they have not always been as aware of the subtle relationship existing between a child's behavior and the events following it as they have been of the behavior and the events preceding it. Traditionally, teaching has emphasized the eliciting rather than the reinforcing events.

When children make appropriate responses, they receive teacher's smiles, words of praise, check marks, stars, and social approval from peers. These are reinforcing events to most children. The precise effect of these reinforcers for an individual child needs further investigation. Many children fail to learn because they do not or cannot respond to the reinforcement provided. Culturally deprived and disturbed children are less responsive to verbal reinforcement than normal children. A child who has been treated inconsistently or has been deprived of middle-class language may not be reinforced by being told that he is "correct." It has been found that in such cases the child is more responsive to reinforcement of a material nature, such as food or tokens.

As in the development of responsiveness to stimuli, there is a developmental sequence in responsiveness to reinforcement. In the beginning of life the infant is comforted or reinforced by food and tactual stimulation. Normally, these reinforcers become paired so that either food or being gentled is reinforcing. These primary reinforcements are associated with talking to the child and thus various words take on reinforcing value. Smiles, tone of voice, physical closeness, and cuddling develop as reinforcers. These events occur most often immediately following the child's desirable behavior. When the child arrives at school, the words associated with these expressions of affection have become conditioned reinforcements. The child's appropriate behavior is therefore supported by words like "correct," "that's right," "good boy," "very good." In school these are paired with check marks, stars, and, later, other symbols or grades so that these now carry the reinforcing value of mother's earlier expressions of love. The next phase is when the child, on completing a task or on behaving well, feels that he is loved even without being told through receiving check marks or words of approval. Thus, internal reinforcement becomes a powerful support during those periods when the child must maintain his attention to a task or pursue a solution to a problem.

At this stage the child is able to employ the project method, although it may be on a very limited scale. He now can complete an assignment and is willing to wait for some time in order to receive a reward such as a grade.

There are higher levels of reinforcement such as status, running fast to be first, or working hard for a position as team captain or class president. There are symbolic reinforcements such as certificates, diplomas, awards, reports, or degrees. Studies have shown that even certain purchasing behaviors of adult shoppers can be maintained and increased by the immediate reinforcement of blue chips or green stamps which are later rewarded by various attractive gifts.

Although the developmental sequence of reinforcement does not appear to be rigidly defined, the primary ones are important and can be used to develop or condition symbolic or token higher-level reinforcers, as well as internal reinforcers. For example, the five-year-old child's behavior problem discussed earlier may have been related to his lack of responsiveness to the reinforcement contingencies in the classroom. In the first step it was desirable to reinforce this child's sitting behavior with M & M candies. By pairing the M & M's with phrases like "good boy," and "that's good," the phrases also became conditioned reinforcers and were used independently of the M & M's to maintain his sitting behavior. Later, when he was working on the pegboard tasks, the phrases were paired with check marks so that these become reinforcers. During this pre-conditioning period, it was effective to present the check marks together with the M & M's and the words of approval in order to strengthen the reinforcing value of the check marks. Later, this pairing was done on a random basis and strengthened the reinforcement of the check marks and words of approval. When the child returned to the classroom, his appropriate behavior and attention were maintained by a schedule of social approval represented by the check marks and occasional words of approval. In this way he moved from his entering reinforcement behavior to the terminal objectives of responding to the reinforcement in a normal setting.

Case Study

Bobby, a four-year-old male microcephalic, attended the university-affiliated nursery school for trainable retarded children. At first his school behavior was uncontrolled and aggressive. Medical reports described him as having a congenital cerebral defect with a significantly small head, mental retardation, and the organic behavior syndrome.

The parents are separated and Bobby lives with his mother. She tends to be hyperactive herself, is very inconsistent in her treatment of Bobby, and appears to resent and, at times, to reject him.

Despite his medical and environmental problems, Bobby responds to both social approval and food reinforcement.

Diagnosis was based on the physician's and the social worker's reports, the teacher's comments, and observation. The school variables were based on information from Prescriptive Teaching (Peter, 1965), as were the recommendations for working with the mother and with the medical and social agencies involved.

Injury. The medical report indicated that Bobby fitted the picture of microcephaly with closed frontanel and a small cranial vault. A skull X-ray confirmed the diagnosis of microcephaly with narrow suture markings. The medical examiner's report stated that he was knock-kneed and flat-footed, and that he walked on the insides of his feet.

Disability. The most relevant disability educationally was Bobby's mental retardation. The second resulting disability was the behavior syndrome. Bobby was extremely hyperactive, reacting especially to noise stimuli; his behavior was quite uncontrollable, unpredictable and often aggressive; his attention span was very short. His coordination was poor, and even the most simple tasks were difficult for him.

Handicap. Because of his retardation, Bobby's adaptive behavior is limited. His verbal ability is limited and consists of a few words used singly.

The mother reported that he had a tantrum every 20 minutes. At school it was observed that this represented approximately the frequency but that these episodes did not occur with the regularity reported. The violent periods lasted about 2 minutes and consisted of waving or shaking his hands, crying out, hitting and biting himself, and when someone was near, hitting, biting, and kicking them. The student-teacher who worked with him was bruised about the arms and legs during the first sessions until she learned how to deal effectively with this behavior.

Most of his behavior in school was aimless in that he did not participate with other children and did not play constructively with toys or equipment. He was destructive and dangerous with equipment. He did not follow verbal instruction nor did he imitate consistently. Sometimes he responded to his name and occasionally came when called. His lack of attention to a task and his distractability resulted in severe learning problems even within the nursery school class for trainable retarded children.

Prescriptive teaching. Judy, an education student, was assigned to work with Bobby for three hours a week for a period of six months. I did some teaching of Bobby for short periods once a week for demonstration purposes. Judy worked in the nursery school giving him individual instruction within the classroom, assisting him in group interaction and in individual instruction in a side room. For the balance of the school time he was retained in the nursery school. The following description, though drastically edited and condensed, attempts to capture the essence of the procedure.

Entering behavior. Bobby's behavior generally was at a two-year level, but he had particular deficits in social behavior. He was subject to the sudden outbursts of violent activity described earlier. His self-care was erratic. He was partially toilet trained. Sometimes he would defecate or urinate when taken to the toilet but he would not indicate when he needed to go. Behavior with peers was very primitive. He pushed them out of his way, hit them with toys, and kicked, bit, or hit them if he did not get something he was trying to take from them. His response to adults was even more immature. He did not look at the adult when his name was called or when attempts were made to show him something. There was little eye contact. Eye fixation on the face, a behavior described in early social development, normally occurs around three months. Bobby's response to other children is approximately at a one-year level. These evaluations helped provide a starting point and the description of normal social development indicated the curriculum for his social learning.

His responsiveness to certain eliciting events not included above are of importance. Although many of his violent outbursts appeared to be unprovoked, he seemed particularly responsive to sounds and a sudden loud noise would bring on an attack. He was distracted by almost any sound. The stimulation of the classroom increased his hyperactivity.

Bobby's response to reinforcement was quite primitive. Initial observation and a few trials showed food to be the strongest reinforcement. Hugging and stroking were effective, but words of approval accompanied by smiles and gestures had little reinforcing value. Tokens (poker chips, beads, etc.) were ineffective.

With this information the assessment of Bobby's entering behavior provided a description of his responsiveness to eliciting events (S), his behavioral development (B), and his responsiveness to reinforcement (R).

Terminal objectives. Because his social behavior was so inadequate in contrast to his motor, physical, and mental development the objectives were primarily concerned with social adaptation such as his ability to:

1. attend to and complete nursery school tasks, including pegboards, form boards, nesting boxes, and puzzles,
2. attend to the teacher during instruction and carry out simple commissions,

3. come when called and respond to his name,
4. indicate toilet needs,
5. maintain non-violent behavior under normal classroom circumstances,
6. cooperate with other children.

The eliciting events and reinforcements supporting the terminal behavior were defined as those occurring in this particular nursery school. The program would be successful if the special intervention of three hours per week of individual instruction achieved these behaviors when Bobby was in class and not receiving special individual instruction. Success would be determined in relation to the number of objectives achieved, the frequency of occurrence of the behaviors, and the proficiency of the attainment.

It is not practical, with such severely handicapped children, to establish rigid criteria for success in defining terminal objectives. The progress that can be expected for Bobby in six months under these circumstances is unknown. The nursery school teacher felt that Bobby could benefit significantly from his school experience if he could develop some of these social skills. It was necessary to establish criteria for success in describing en route objectives in order to have a basis for deciding when to move to the next step. If the terminal objective was a change of placement for Bobby, then criteria for that placement should be stated, but in the nursery school concern was for the development of more adaptive social behavior so that he could profit from the school experience. A pass-or-fail concept was inappropriate in this case.

Developmental curriculum. The lowest level social skill to be acquired was to attend to an adult when the adult spoke to him. The highest levels were the use of language and the practice of impulse control. The abilities in between are described as a sequence of developmental tasks. During the teaching process, the methods of achieving these developmental abilities were determined through trial and error or a cut-and-fit procedure. The objectives were for him to:

A. attend to an adult and respond to his name;
B. come when called;
C. sit in a chair on command;

D. respond to the word "no";

E. assemble pegboards, form boards, and puzzles;

F. engage in parallel play;

G. integrate his behavior with others;

H. take turns with others;

I. help others;

J. indicate toilet needs;

K. carry out a simple commission;

L. recognize common objects;

M. recognize people and their names;

N. speak in sentences;

O. maintain non-violent behavior.

Teaching. This description of the teaching process is limited to the main events used to bring about the progressive changes in behavior described above. The consistent approach Judy employed toward Bobby was firm-kindness (Peter, 1965). The teaching process is described by (a) stating the en route behavioral objectives, and (1) explaining the use of specific new instructional materials and methods employed as the eliciting events, (2) recording Bobby's behavioral response, (3) defining the principal reinforcements employed, and (4) describing concomitant developmental changes in behavior as they occurred in relation to the responses directly controlled by the teaching process.

(a) En route objective. When called by name, Bobby will look in the direction of the speaker 8 out of 10 times. (80% is adequate since it is not always necessary for the child to look at the speaker, although it is a desirable response at this stage.)

(1) Eliciting event. Bobby was seated in a small chair facing Judy and in a manner barring his escape. She said, "Bobby," and waited for him to look in her direction. In the beginning when he would stare to one side she would say, "Bobby, look at me," and gently turn his face in her direction.

(2) Behavior. Looking in the direction of the speaker was rewarded in the beginning and this was shaped until eye contact was the only reinforced response.

(3) Reinforcement. Looking in the appropriate direction was immediately reinforced by simultaneously putting an M & M candy in his mouth, saying, "good," and smiling. When he learned to look at the speaker's face with only the help of "Bobby," or "Bobby, look at me," the M & M's were given at random to strengthen the reinforcement of the smiles and "good." At the end of the first hour, saying "Bobby" elicited the response at least eight out of ten times. This was reviewed at the beginning of the next and successive sessions.

(b) En route objective. Bobby will come when called, consistently ten times.

(1) Eliciting event. At first, Judy held Bobby by both hands while he stood facing her. She said, "Bobby come!" or "Bobby, come here!" and then pulled him toward her. Later she held him by one hand, and still later she called him from a distance of six feet. When this was achieved, she moved him from the small room where they were working to the big classroom with the other children.

(2) Behavior. In the beginning he struggled, but when he came without resistance one hand was released and he was only guided toward Judy. When he came without any physical guidance or pulling to initiate the behavior, both hands were released. When he came consistently 10 times while in the small room, he was moved to the classroom where distractions reduced success to about 50%. Many trials at various distances were needed before the criteria of 10 consistent performances occurred.

(3) Reinforcement. When Bobby came to the teacher, he was first given an M & M, "good," or "good boy," and a pat on the shoulder.

(4) Concomitant development. The nursery school teacher reported general improvement in his behavior including coming when called and some reduction in violent behavior.

(c) En route objective. When told to sit in a chair, Bobby will follow the instruction and will sit for two minutes without restraint.

(1) Eliciting event. Bobby was told to "sit" or "sit down" and was then gently guided or pushed into a sitting position in the chair. After four presentations of this event he was simply told to "sit" and in the class situation the chair was indicated by pointing.

(2) Behavior. He sat without resistance and quickly learned to follow directions regarding which chair to sit in. Remaining seated occurred infrequently.

(3) Reinforcement. Sitting was immediately followed by "good," a smile, and a pat on the shoulder. These reinforcements singly and in combination were continued at intervals to maintain the sitting behavior. If he got up before the two minutes expired, Judy turned away from him.

(4) Concomitant development. The nursery school teacher reported that Bobby had altered his running away behavior. Previously, he ran away when approached, but now he usually would wait when being approached by her or other adults working as volunteers. He usually came when called and sat at the table when instructed to do so but he did not wait while the fruit or juice was being served. He was out of his seat frequently and usually interfered with other children.

(d) En route objective. He will stop his present behavior on the command "no!"

(1) Eliciting event. Bobby was instructed to sit in a chair and to "stay" or "stay in the chair." If he stood up or, preferably, at the initiation of standing up, the word "no" was said emphatically and paired with restraining him from completing the act of standing. After repeated pairings, the word "no" was effective by itself. When "no" became effective in this situation it was used similarly as the event to elicit a cessation of other behaviors such as hitting children or the teacher.

(2) Behavior. Saying the word "no" became an effective means of stopping his undesirable behavior. When "no" accompanied the initiation of an act he returned to the appropriate behavior. When it followed the completion of the act he stopped but did not return to the appropriate behavior.

(3) Reinforcement. When, under direct close observation, it was possible to associate the word "no" with the initiation of an act and when "no" stopped the continuance or completion of the act, Bobby was given praise and a smile. If he persisted after "no" the word "no" was repeated sharply and was paired with a slap, usually administered to his wrist or to the back of his hand. In the case of kicking, when the initiation of the behavior was observed, Judy was frequently able to intercept his shin

with the side of her foot so that as he tried to kick he was blocked, with somewhat painful consequences, and "no" said sharply was paired with this. In this way he learned to respond to "no" in its two meanings; one, as a command to stop and the other as an aversive reinforcement after doing something wrong.

(4) Concomitant development. The immediate improvement noted in the classroom was the ability to control his undesirable behavior from a distance. Also in the classroom other language became effective. In "No, Bobby, stop that!" the word "stop" took on behavioral meaning. Probably the most significant change was the emergence of self-control. He would give evidence of initiation of violent behavior and then stop himself. From this point on, learning of desirable social behavior was accelerated.

(e) En route objective. When presented with a variety of pegboards, form boards, and puzzles, Bobby will assemble them correctly.

(1) Eliciting event. Starting with the presentation and demonstration of simple pegboards and progressing to more complex pegboards, form boards, and preschool jigsaw puzzles, Bobby was given materials that involved moving from simple to complex motor and perceptual tasks and, with the puzzles and games, additional cognitive stimulation was given him. The earlier tasks were short and the later ones required longer periods of time.

(2) Behavior. Now Bobby came when called, sat in a chair, attended to teacher. His violent behavior was under some degree of stimulus control. He performed these new motor and perceptual tasks without excessive irrelevant behavior. He was distracted frequently but returned to the task at hand when instructed to do so.

(3) Reinforcement. At first he was reinforced with praise for each peg placed in the board, then only for a row of pegs, and finally only for completion of a board or puzzle.

(4) Concomitant development. He participated in many of the same activities as the other children and even excelled in some things. His irrelevant hand movements decreased and he cooperated more in dressing himself.

(f) En route objective. When introduced to a parallel play situation, Bobby will join in and participate along with the group.

(1) Eliciting event. Starting with a row of children holding a rope, Bobby was guided to hold the rope. From this simple beginning, the group progressed to following the leader while holding the rope, to following while holding hands, to following without holding, to following through various obstacle courses, to marching and rhythm band. By this sequential arrangement of events, a progression of parallel social behaviors was elicited.

(2) Behavior. Bobby responded at first by participating for only short periods and dropping out when distracted. When he would get out of line a warning "no" would usually be effective in causing him to stop. Sometimes he would need guiding to return. He progressed through the stages of this developmental sequence with the other children.

(3) Reinforcement. He received praise for parallel play at first but modeling behavior developed and he showed much evidence of pleasure when he was successfully engaged in these activities.

(4) Concomitant development. Activities were now motivating and he was eager to enter into them. He engaged actively in getting out equipment for some of these games. Earlier reinforcements could now be used as motivators. If shown an M & M, he would work diligently at the assigned task to gain the reward.

(g) En route objective. He will participate with the class in activities such as getting out and putting away equipment.

This behavior appeared spontaneously. It was only reinforced as was that of the other children and it appeared to be self-rewarding or modeling behavior.

(h) En route objective. When required to take turns with another child, or within a group, he will wait his turn and then participate.

(1) Eliciting event. The first material employed was a large set of nesting boxes. After Judy and I demonstrated the procedure of putting one box inside another, alternate boxes were given to Bobby and another child. Each was restrained until the other took his turn.

(2) Behavior. After they were guided through the procedure they were able to wait and take turns.

(3) Reinforcement. Both boys were rewarded with food and praise for the correct completion of the project. After several successful trials, the only reward was completion of the task itself.

(4) Concomitant development. Generalization of this skill was accomplished with little further instruction. Bobby, for the first time, was willing to wait his turn in games and in some other class activities.

(i) En route objective. He will help others.

The objective of helping others was accomplished with little instruction. Spontaneously he tried to help move the slide and chairs, to help other children with their coats and buttons and he initiated other helpful behaviors. These offers of help were reinforced when observed.

(j) En route objective. He will indicate when he needs to go to the toilet.

(1) Eliciting event. Bobby had been taken frequently to the toilet and was reinforced when he defecated or urinated in the toilet. Whenever he looked as though he needed to go, he was taken to the bathroom door and asked if he needed to go.

(2) Behavior. He went to the bathroom door when he needed to go.

(3) Reinforcement. In the beginning reinforcement, in the form of praise and occasionally food or a small plastic toy, was given both for going to the door and for urinating or defecating. Later reinforcement was provided only for completing the act.

(4) Concomitant development. Toilet training and cleanliness improved throughout the day as well as in school.

(k) En route objective. When given a commission he will carry it through to completion.

(1) Eliciting event. He was given instructions, such as, "Bobby, take this ball and put it in the box and then come back here." He was then handed the ball.

(2) Behavior. At first he put the ball in the box but it was necessary to call him back. When he carried out a commission without intermediate instruction, it met the criteria for success.

(3) Reinforcement. Praise was awarded only for successful completion of the commission.

(4) Concomitant development. The ability transferred readily to other situation .

(l) En route objective. He will identify common objects by name.

(1) Eliciting event. Common objects such as balls, blocks, spoons, forks, scissors, dishes, and shoes are placed on the table. Bobby is instructed to "Show me the ball," "Touch the scissors," "Give me the spoon," etc.

(2) Behavior. He usually identified items correctly. When he could not identify an item, Judy picked it up and said, "This is a spoon."

(3) Reinforcement. Verbal approval consisted of "Right, this is a spoon."

(m) En route objective. He will identify people by name.

(1) Eliciting event. I put a ball in his hand and said, "Bobby has the ball." I took the ball and said, "Give the ball to Peter." I then repeated this with Judy and the other student teachers, "Give the ball to Judy," "Judy, give the ball to Bobby," and so forth.

(2) Behavior. This was continued with various combinations of persons until he could identify all the adults and children in the classroom.

(3) Reinforcement. The adults thanked him when he delivered something to them.

(4) Concomitant development. He was now able to respond to much more conversation. He could identify persons and go on errands within the class. He called most of the children by name.

(n) En route objective. He will use meaningful speech and speak in sentences.

This was encouraged throughout the program and in particular through tasks (k), (l), and (m), where he was asked to repeat instructions. He spoke quite distinctly and sometimes spontaneously used short sentences. When he was able to express a need, it was frequently fulfilled so that speech was rewarded. The volunteers and visitors were impressed with his language ability so he received additional attention for speech.

(o) En route objective. He will be able to maintain non-violent behavior.

He achieved this to the degree that his behavior was acceptable and that he could be controlled by a warning signal such as "Bobby, stop that!"

Bobby was described by visitors to the classroom as a highly motivated boy who was trying hard to be a good boy and to please others.

Conclusion. The terminal objectives were achieved to a high degree within the six months of the special intervention. It was obvious that he had changed as a stimulus object. At first he was regarded by most visitors as a bad boy and was now seen as a good boy. He had learned a wide repertoire of adaptive cooperative social behaviors which, through the developmental approach, were generalized or transferred to situations outside the formal, or classroom, situation.

Some details which were omitted were trials that proved ineffective. The procedure in the beginning was so highly specific that success was highly probable for each step. Each behavior was established before the next was elicited. Each session reviewed the learned behaviors before the new one was attempted. If the new eliciting event did not stimulate the desired response, something else was tried. If, at the end of a session, the new behavior had not been elicited, the teacher returned to the last successful elicitor so that the session ended with Bobby receiving reinforcement. Each session concluded with his being successful. Judy spent the last few minutes playing simple pat-a-cake games and holding and hugging him before he left.

This case includes many of the early stages of development of cooperative social behavior. Only a few seriously impaired children require this degree of structured sequential teaching. When incidental learning is seriously impaired in cases like this or in cases of infantile autism, the teacher must be responsible for each task. In cases where incidental learning is less impaired the teacher may have to deal with only a few behaviors and the child will then develop new adaptive behaviors through experience in the natural environment.

Bobby's case illustrates a child development approach including some of the early individual behaviors required for successful cooperative behaviors. It also illustrates classic and operant conditioning as well as modeling behavior.

Modeling, in the development of social behavior, is an important aspect of learning which should be remembered by all teachers. Children pattern much of their behavior after that of adults. Teachers who in their own behavior illustrate consistent cooperative social behavior will facilitate the development of these behaviors in children.

Small Group and Classroom Applications

The learning principles described in this paper have been employed in many classrooms for disturbed children. The learning situations are structured so that cooperative social behaviors are both elicited and reinforced. In one intermediate special class of 15 children with behavior and learning disorders, the class was so badly behaved that most of the teacher's time and effort was expended in an attempt to maintain order. An experimenter told the children that for every minute the class maintained order and followed a few simple rules of good conduct, a poker chip would be placed in a glass that was in full view of the class. The entire class would be rewarded by special outings, parties, or other festivities when the required points were earned. Within a few days the class was working for two hours without interruption. By retaining the reinforcement system, the regular teacher was able to maintain the improved behavior. In this situation the children gave up encouraging and reinforcing the disruptive behavior of their classmates. Instead they supported the more appropriate social behavior that facilitated study and learning.

The teacher in a primary class for educationally handicapped children reinforced each individual child by dropping an M & M candy in a glass. Later all the M & M's earned by all the children were shared equally. It appeared that all the children wanted each child to be successful, resulting in a high level of cooperative social behavior.

In Prescriptive Teaching classes we start at the child's developmental level and structure the experiences so that he achieves progressively higher levels of cooperative social behavior. The application of Prescriptive

Teaching to the classroom for children with learning disorders or emotional disturbance results in organizing a highly structured classroom which quickly brings the children's behavior under control of the educational process. The children work individually on their sequentially organized assignments. After considerable success in this phase, they are involved in parallel activities, such as all of them listening to a story being read to the class. The next phase involves taking turns in answering questions. When this level of cooperative behavior is well established, some of the children can be assigned to working in pairs on projects requiring limited social interaction. Children in these classes must not be moved too rapidly into social involvement. Much success at each level is required before moving to the next step.

Summary

The structure of the process of education presented in this paper provides us with appropriate questions to ask when a child does not respond adequately. Therapeutic teaching should begin by asking:

1. What is the child's present functioning or entering behavior?

   Evaluate his response to the present environment and find what his behavioral response is when we vary the eliciting and reinforcing events.

2. What are the terminal objectives?

   If the child has the capacity for normal development, then the terminal objective is stated in terms of the regular classroom environment, the regular eliciting events and instructional methods, the appropriate behavioral response, and the regular reinforcement contingencies.

   In the case of autistic, schizophrenic, trainable mentally retarded, and others with severe organic defects where a wide range of adaptive behavior is unlikely, the objective is in terms of the child's response to some normal eliciting and reinforcing events so that he can learn maximally in the normal environment. In these cases it is not possible for the child to achieve full developmental responsiveness in the normal environment.

3. What are the sequential steps in the developmental curriculum?

   Educational literature, instructional material, and child development studies provide guidelines for finding these answers.

4. What sequence of eliciting events should be presented in order to stimulate these progressive developmental changes?

Our whole instructional technology, educational materials, toys, games, etc. provide stimuli with a high probability of eliciting these behavioral responses.

5. What reinforcing events should be employed to strengthen these behaviors when they occur?

Observation and testing situations can provide opportunities to discover the child's present stage of reinforcement development. The effective reinforcements are used to initiate the learning process. Through pairing or conditioning the reinforcements, the child is brought to the developmental level where he responds to normal contingencies.

In answering these questions, events outside the teaching-learning situation influence our answers but it is these answers that provide the therapeutic teaching program. The structure of the process of education can be used in dealing with any learning problem whether it be in the areas of academic learning disability, emotional disturbance, or social behavior problems.

References

Peter, Laurence J. Prescriptive Teaching. New York: McGraw-Hill, 1965.

CHAPTER X OPERANT REINFORCEMENT OF INFANTILE AUTISM

C. B. Ferster

I will discuss some of the ways in which the behavior of the autistic child is related to his social and physical environment (Ferster, 1961, 1965). In doing this, I will focus on the extreme deficits of infantile autism (Kanner, 1944, 1943; Bender 1953); but the general approach and the special techniques that I may suggest for dealing with this problem are intended to apply to less serious forms of childhood schizophrenia and even to adult psychosis.

The autistic child, often three years of age or perhaps younger, spends long periods of his day without activity--periods in which he may lie on the floor, staring vacantly into space, or in which he sits in one position, barely moving. Whe he *is* active, his activity is of a very simple sort--he moves his fingers within his visual field, he walks up and down, he touches spots on the wall again and again, he rubs his hands together, flips a foot repeatedly, or swings a door back and forth. These performances have little effect on his physical or social environment, nor are they peculiar to the autistic child. Almost all children, and many adults, show lapses of activity and repetitions of simple movements; but

Reprinted from An Evaluation of the Results of the Psychotherapies by Stanley Lesse, M.D., MED.Sc.D., 1968, Charles C. Thomas, Publisher, Springfield, Illinois, U.S.A.

Note: Support for the research at the Linwood Children's Center from the Aaron E. Norman Fund, Inc. is gratefully acknoweldged.

for the autistic child these have a high frequency relative to other ways of behaving.

Most autistic children engage in some degree of social behavior, but such behavior is restricted largely to actions which affect others aversively. He stands in front of the door and screams until someone opens it for him. He prevents the removal of a shoelace he has been chewing all week by screaming when anyone attempts to take it. Atavisms, such as temper tantrums, may be so unpleasant that one hesitates to interrupt a child's compulsive or ritualistic activity for fear of bringing these on. Such behavior occurs, of course, in children who are not psychotic, but in the autistic child these primitive forms of social control are especially prominent. Because they have such a consistent, direct, and large magnitude social effect, they are highly resistant to change.

The autistic child lacks many of the behaviors of the normal child. He is frequently mute, even though he has nonvocal behavior such as tugging on a sleeve or making some effective gesture. Sometimes the vocal behavior is only apparently verbal. It may be only imitative or echoic, or it may be an intraverbal chain that has no real environmental reference. Even fairly extensive vocalizing in a less-damaged schizophrenic child may lack any functional connection with a listener. The child may speak at length, and coherently, but in a monologue; a point-to-point, conversational relation will be lacking. Except in the form of a demand, truly verbal behavior may be absent. The child says, "cookie" in the presence of one who has given him cookies in the past, but he will not say "cookie" simply in response to the question, "What is that?" He will say, "Out!" when he is in front of a locked door but not in response to the question, "Where's Johnny?"

A normal child's daily activity is under the control of myriad environmental changes. He is influenced by the facial expressions and voices of his parents, but the details of his toys and books, by the coarse or fine features of his clothing, and so on. We tend to notice those parts of the environment which make a difference for the reinforcement of our behavior. The autistic child's minimal repertoire requires few distinctions for it to be effectively reinforced. Even though the child is in an environment that is potentially reactive to his behavior, he will not be controlled unless he emits behavior which is differentially reinforced. Experimental measurement of what an autistic child in fact sees and hears frequently shows major deficiencies.

The autistic child lacks among chains or sequences of behavior in which each performance is a necessary condition for the next. This deficiency follows directly from the absence of the simpler behaviors. Socially derived reinforcers are impossible without large amounts of behavior under close control by people and things. Reinforcers such as attention, approval, praise, money, power, progress, or achievement comes from a complex history.

In summary, the autistic child has little potential for either changing or being changed by a normally reactive milieu. The autistic and normal child differ, first, in the relative frequency of the various kinds of behavior. Almost any adult or child will, from time to time, show lapses in behavior or simple repetitive acts like tugging on one's ear or mustache, but for the autistic child these are the major part of his activity. Second, the autistic child completely lacks many kinds of performances by which the normal child interacts with his environment.

The autistic child's deficits could come from his organic structure or as a result of a history of experiences or a combination of both. No one has yet identified any plausible organic factor for autism while, on the other hand, it is not difficult to imagine environments which could produce behavioral deficiencies as great as those due to brain damage. Therefore, I assume, as a working hypothesis, that special environments at one or more stages of growth are responsible for the autistic child's deficient repertoire.

If the autistic child is organically intact, it behooves us to examine the ways in which the child and his environment may interact so as to strengthen or weaken his behavior. The principles of operant reinforcement provide a framework for describing how the child's social milieu might have produced the gross deficiencies we see.

Reinforcement and extinction are the major ways of creating and maintaining behavior, and weakening or removing it. Many factors in the natural environment may produce nonreinforcement of a child's behavior and hence are potential conditions for weakening or removing parts of the child's repertoire (e.g., parental inattention, setting of limits, or competition for the parent's behavior by other persons). Social situations inherently lead to large amounts of extinction. The mother may not answer the child because she is on the telephone, immersed in a book, or otherwise occupied.

Many other changes in the environment are more severe, however, and lead to situations where large segments of the child's behavior will not be reinforced. A psychotically depressed parent, a drug addict, or an alcoholic will react to a child's behavior infrequently. With some parents, *apparently* favorable environments frequently turn out to have no reactivity to the child. I observed a parent, for example, who talked to her child all day long much in the manner of the average mother. When I examined the relation between the mother and child more exactly, however, there were very few instances where the mother's behavior occurred as a result of the child's except when she explicitly prevented him from having an effect on the physical and social environment. At one point she apparently read a story to the boy, but he gave no indication of paying attention and had to be captured first. Each time the *boy* acted on the physical environment, however, the mother almost magically prevented the effectiveness of his behavior. He reached for a lamp, she stopped him; he tried to go out, she prevented the opening of the door; he started to play with a younger brother, she intervened. What initially appeared to be a normally reactive environment turned out to be consistently nonreactive. It would have been difficult, at first look, to have imagined this mother capable of such consistent effort.

Much extinction of a child's behavior occurs indirectly because the environment is not uniformly reactive. The kind of behavior that is reinforced by a listener while he is smiling, for example, is very different from the behavior he will reinforce when he is frowning. A child's behavior conforms to the environment because it is reinforced on one occasion and extinguished (weakened) on another. The essential ingredients of the interaction are (a) the child tends to ask for cookies because asking for things gets them, (b) the verbal response is effective only in the presence of a parent and, (c) the reinforcement *only* with the presence of the parent reduces the frequency of the response in the absence of a person. Setting limits for a child is just this kind of specification of reinforcement contingencies. A limit specifies an occasion on which a response is not reinforced (or even punished).

When a child's behavior is narrowly under the control of limited parts of the environment, it may be severely weakened. Should the environment suddenly change, the child would be under the control of circumstances where much of his behavior has been extinguished. The narrower the range of stimuli which occasions when the child's behavior may be reinforced, the more fragile the behavior is to disruption because so much nonreinforcement has occurred in narrowing the control by the environment. Yet, as the child matures, the community must progressively restrict the occasions on which the child's behavior may be emitted. A given community may reinforce baby talk at one year of age, but it will not at the next year. Between the normal community practice and extremely narrow environmental control lies a dangerous potential for weakening behavior.

How much a sudden change in a child's environment will disrupt its repertoire will depend upon how much behavior the child has under the control of how many features of the environment. Certainly a radical change in the environment will disrupt any child. But the extent of the disturbance will be much greater if the child's history includes experiences in which a very wide range of stimuli comprised the occasions on which the child's behavior was not reinforced. Such would be the case if a parental environment reinforced responses very narrowly under very special circumstances. Parental practices could result in such narrow environmental control that a sudden shift in the environment would leave the child literally denuded of his behavior.

Consider, for example, a mother who did not interact with her child and yet remained in the close vicinity so that all of the child's behavior in respect to her went unreinforced at the same time that a nanny met all of the child's needs. Such a mother is perhaps hard to imagine, but I assure you that she exists. When the nanny left, after a year of mediating nearly all of the child's reinforcers, the child progressively lost all of her behavior, until after two months she became incontinent, mute, and had to be withdrawn from a nursery school with a classical diagnosis of autism.

The child had acted on the environment only through the young nursemaid, whose absence removed the support for all of the child's behavior. The nursemaid's sudden absence would not have had such dramatic results had not the mother's nonreactivity to the child led to consistent nonreinforcement of the child's behavior al' year while the nursemaid was in charge. Had the mother instituted a sensitive interaction with the child when the nursemaid left, the child's behavior might have been reinstated on the various occasions where it had been extinguished.

Punishment and aversive stimuli also weaken behavior because of nonreinforcement rather than direct effect. While we frequently use electric shock in the laboratory as a model aversive stimulus, the major techniques of aversive control on the normal human ecology are the withdrawal of positive reinforcement (and hence extinction), as in the case when a child is incarcerated in his room, in a cage, or in a closet; as in criticism, anger or other kinds of disapproval when the parent is not inclined to reinforce the child's behavior, or even as in corporeal punishment where the most profound effect may come from the parent's low disposition to positively reinforce rather than from the noxious effect of the punishment. Punishment does not always have its intended effect, and its very ineffectiveness may, with some parents, simply increase the frequency of its application. Under the circumstances, punishment may serve as a major way that a child's behavior is weakened, much as in other methods of extinction.

The circumstances surrounding and particularly preceding punishment may have general suppressive and emotional effects as indirect by-products. When the child is punished a great deal, virtually every part of his environment may generate anxiety. The child's own behavior is also a part of his environment and is selectively correlated with punishment. Only those behaviors which actually alter the external environment are likely to be punished. Performances which have little effect beyond the child himself--staring out of a window or moving the fingers in front of the eyes--are less likely to be punished. Hence, any productive behavior may be anxiety provoking.

A parent might also weaken a child's behavior by reinforcing primitive behaviors so strongly that they preempt normal growth and development. Many parents, unusually sensitive to temper tantrums, crying, or injuries strongly reinforce them. The atavisms become progressively more severe and eversive as the child has a greater influence on the parent with more intense forms. Sometimes a parent reinforces crying when the child is sick, and it becomes difficult to eliminate this crying when the child is well. The parental practices often lead to a strong repertoire, difficult to extinguish, which has many of the properties of the compulsive gambler. To appreciate the amount of reinforcement given to atavistic behavior, one need only compare the behavior of a child who sits quietly playing with his blocks with that of a child who suddenly screams, throws toys about, kicks, and soils himself. At least on the first occasion, the child who is playing will be ignored; the major attention will be paid to the child who *raises the ruckus*. Atavisms bother people; these behaviors have an immediate emotional impact, and they generate behavior in others to end the annoyance.

Self-destructive behavior of the autistic child generates another kind of aversive effect on those around him. It is almost impossible not to react vigorously, emotionally, and in depth to a child who is forcefully striking a cement wall with his head or is picking his flesh so that bones are exposed. The work of Lovaas, et al. (1965) shows some of the ways in which self-destructive behaviors are maintained in the autistic child's repertoire because they control the behavior of those in the child's milieu.

No *one* of these ways of weakening behavior is, of itself, likely to be responsible for the extreme deficiencies of the autistic child. The total failure or loss of an autistic child's behavior would logically come from a combination of circumstances, whose common influence is that they weaken the child's operant repertoire. Several general characteristics of a young child make him susceptible to extreme loss of behavior. There is the limited physical and social mobility of young children. Prior to school age, most possible ways a child can change environment are mediated by one parent.

It is possible to interfere with the behavior of a child of any age as, for example, did the mother of one autistic child who kept a bottle in his mouth for two or three hours at a time no matter how much the child cried, screamed, or struggled; this mother also fed the child from a baby food jar with a one-gram spatula because the sight of food running down the baby's mouth distressed her. Such a parent is perhaps unusual, even as unusual parents go since a parent is usually more disposed to interfere with a child's behavior when he is tugging on her skirt, breaking china, interacting with other children, or dirtying furniture or the kitchen floor than when the child is in the crib. In other words, the more effective the child is in producing major changes in the social and physical environment, the more occasions he presents to the parent to interfere with his behavior. It is paradoxical that a parent who is most disposed to interact with a child is the one who can potentially do the most harm.

Some autistic children have grown up in rural environments where one parent, exclusively, maintains all of the child's behavior. An urban or suburban community, however, does not guarantee social contact with neighbors and lacks the characteristics of a rural environment, which may support much behavior through its physical properties if the parent allows the child to engage it. Occasionally an entire family conspires to ostracize a child. The father, frequently absent from the house, either ignores or punishes the child. The sibling, to the extent that he deals with the child at all, duplicates the mother's practices.

Many different kinds of parents may interfere with the child's behavior in many different ways: the economic status of the parents may be high or low; they may be psychotically incompetent; or they may be highly effective people. The actual practices of the parental environment give us the most information about how the parent produces the autistic child's repertoire. The common denominator among the parents of autistic children lies in their manipulation of the child's milieu in such a way that the child's behavior is weakened.

The literature is accumulating reports of new ways of experimentally arranging environments which have new magnitudes of effect toward building new behavior for autistic children. These experiments all confirm the importance of the environmental history of the autistic child. My experiments in Indiana (Ferster and DeMeyer, 1961, 1962) developed procedures which conformed the child's behavior to a carefully reactive environment in small steps, like the gradual processes of normal growth and development. The children in the experiments identified letters of the alphabet, withheld their behavior on inappropriate occasions, worked in the experimental environment for delayed reinforcers such as swimming or playing outside, and performed other kinds of complex behavior. In some experiments, we maintained productive behavior regularly for two hours and more with children about whom it might have been said, "You can't keep their attention for more than five minutes."

The very impressive case study by Wolf, Risely, and Mees (1964) used the same kind of functional analysis therapeutically in the *total* hospital milieu instead of in an arbitrary experimental procedure as I did. Their research had the therapeutic goal of decreasing the frequency of primitive behaviors and building new productive ones. Lovaas' (1964, 1965, in press) experiments demonstrated techniques for creating and measuring broad areas of social behavior using food initially and shifting thereafter to normal social control. These experiments, as well as those of Baer and Sherman (in press), Bijou (1954), Davison (1964), Sherman (in press), and others have demonstrated that we are limited only by our ability to arrange effective environments.

The natural milieu lends itself to the same kind of functional, objective analysis of behavior that has been successful in the experimental laboratory. The environment, on the one hand, and the behavior of the child, on the other, can be as simply described by objective description as can any other natural event in the natural world. We can also describe the way in which the environment and the child's behavior precede and follow each other: *On what occasion did the behavior occur and what were the actual consequences in the environment of the behavior? Was the parent smiling or frowning when the child asked for a cookie? Did he get the cookie?* These data have the same potential objectivity of other events in biology. Our difficulties arise mainly when we try to explain *why* the behavior occurred.

Identical performances may have such different relations to their controlling environment that they are effectively different behaviors. The child who says "cookie" in the occasion of the spoken word "cookie" is engaging in behavior which is only partially under the control of a cookie which is present, or in response to, "What is in the jar?" A cry elicited by falling may be very different from crying which occurs when the child is put to bed at night. The first may be a reflex; the second may be an operant maintained because the parent comes. A functional analysis of behavior in the complex milieu is generally difficult because there is not enough information about the functional relation between the behavior and the controlling environment. With enough facts, however, the descriptions may be quite exact.

A clinical setting as a laboratory has an advantage in that there is a host of durable reinforcers in the very fabric of the child's life with which to reach therapeutic goals. It has the disadvantage in that the environment may not be arranged as consistently as is possible in the laboratory experiment. Yet there remain the problems of a functional analysis of a complex milieu and the discovery of ways in which natural science methods may be applied to the development of therapies.

I am currently engaged in a clinical project at the Linwood Children's Center, a residential and day-treatment center for non-brain-damaged autistic children in the Washington area. Jeanne Simons, the director of the Center, and I began collaborating in a clinical project when we discovered that our methods and procedures had much in common. In a sense I am a sheep in wolf's clothing, and Jeanne Simons is a wolf in sheep's clothing. Even though my background is in the animal laboratory, my customary approach to an experiment is essentially clinical.

Procedures in my animal and human experiments are carefully designed to meet the repertoire of each individual subject; there is a day-by-day interaction with the experiment in which each procedure derives from the results of previous procedures. Conversely, when I observe my coworker carefully, I discover that she manipulates the environment in the clinic much as I do in the laboratory. I use relays and automatic equipment for arranging contingencies in the laboratory; she places limits on the child's behavior, gives or withholds food, attention, automobile rides, and toys. Both of us are controlled by how our procedures alter the behavior of the child. We evolved a natural collaboration which is, in some sense, an attempt to produce a model for collaborative work between the clinician and an experimental psychologist. I knew many ways in general of modifying behavior; I could identify many behavioral processes by direct observation of the critical interactions between a child and his milieu, and I could suggest many processes and procedures which could be applied to the therapeutic goals of Linwood. I lacked, however, the wide range of skills in dealing with the child's total life which an experienced therapist develops over a period of years. Though I might suggest procedures which could be essentially correct, they could fail or not have their maximum success because I failed to adjust a host of incidental details which might be critical. Just as my lack of experience in therapy with children made it difficult for me to extend some principles of behavior to clinical situations, the current clinical procedures in use were difficult to describe in simple, objective terms.

We particularly would like to know what particular part of clinical procedures are effective and on what their effectiveness depends. As I observed the interactions with the children directly, however, and from the vantage point of an experimental analysis of behavior, I was able to sort out some simple descriptions of the critical dimensions of the interactions from a complex social environment. The power of a functional analysis lay in the objectivity of the descriptive terms. What the child did was simply described. What Jeanne Simons actually did was equally objective. The following are some excerpts from my notes on Jeanne Simons' procedures at Linwood. The first excerpt describes a straightforward application of differential reinforcement and extinction. The application depended upon correct timing and other kinds of progress the boy had made.

> Andy is a husky nine-year-old boy who has a lot of verbal behavior, most of which is echoic. When he came to Linwood, he beat his chest with his hands until he became black and blue and repeated everything that was said to him over and over again, particularly when it was an instruction, a reprimand, or a request. He reacted similarly with remarks of the same form made to other children. His voice had a peculiar guttural quality made by tightening up the voice apparatus in the back part of his mouth.
>
> Andy provides an interesting case of the separate functioning of different kinds of speech. The fact that he echoed virtually anything said to him did not mean that the same verbal forms are an effective part of his repertoire in other forms of verbal function. At supper Jeanne Simons began extinguishing the echoic behavior. The supper provided a very good occasion because the food was an effective reinforcer. Each time something was needed--like more meat, milk, or

> dessert--Andy would beat his chest and vocalize gutturally, or he would imitate one of the other children's requests. If he asked for something and was told to wait a minute, he would repeat, "Wait a minute--wait a minute." Jeanne reduced the amount of vocal interaction with the other children to the point where he had very little to echo. She spoke to Andy very little and in a whisper. As she withheld consequences (e.g., more milk) when he beat his chest and spoke with his unnatural, guttural tonal quality, he said one time in a normal voice, "I want milk," instead of the usual, "Andy wants milk." The result was very dramatic within the course of three or four supper hours. The frequency of the chest beating declined dramatically in spite of its persistence for several years.
>
> The reinforcement and extinction took place in the context of the normal supper hour in the midst of the usual interactions with eight other children. The episode occurred when Jeanne Simons judged that many other interactions she had with Andy made it likely that a normal vocalization was possible and that she could handle any of the complications which might arise.

The following account describes the first sustained interaction with an autistic girl who had just come to Linwood. The account illustrates that the therapeutic procedures consist of very active manipulation of the child's environment. In the space of approximately thirty minutes, there were perhaps one or two hundred instances where what Miss Simons did was explicitly contingent on what the child did. Furthermore, all of the contingencies were paced toward a therapeutic goal by requiring slightly more from the child as each reinforcement changed the child's behavior cumulatively.

> Kathy is a mute autistic girl who interacted with her environment very little at home or at Linwood, except to cry continuously but softly, in contrast to the extreme magnitudes so often encountered in autistic children. On this day, for the first time, she was placed on a rocking horse while Jeanne Simons sang and rocked her. The rocking horse and singing stopped the crying, and Jeanne then began introducing brief periods while she kept on singing but not rocking. They seemed designed to fade the singing and to have the girl rock herself without crying. There was a very careful sensing of how long the period of no-rocking could last. Later Jeanne stopped the rocking when the child cried. The toy which the child compulsively carried was placed on a nearby table, and the table

was moved very close to the girl so she could take it back quickly. This all occurred within perhaps ten seconds and was the first time in which Jeanne changed Kathy's environment somewhat and required some behavior of her.

Jeanne was continuously in action, moving from one place to the next vocalizing at the appropriate times, moving the girl and rocking the horse. When Kathy took the doll off the table, she rocked *herself* slightly. Then, Jeanne sang only when the *girl* rocked, and Kathy rocked the horse by herself more frequently. Jeanne rocked the horse with increased vigor in the manner of a probe. She took the doll out of the child's hand, placed it on the table again, but allowed her to take it back soon. She continued to sing only when the child rocked and then replaced the doll on the table and the girl very calmly took it back. At this point, the girl, herself, placed the doll on the table. One would guess that under almost any other circumstances, other than this very subtle interaction, removing the doll from the girl would produce screaming. Then Jeanne rocked the horse with a magnitude greater than heretofore, perhaps as another test or a probe of this sudden activity on the child. She placed the child's feet on the footrest of the horse, but Kathy didn't keep them there. The intensity of her voice as she sang was paced with the intensity of the rocking, and the toy still remained on the table. Here, for the first time, the child had been without the toy doll for over a minute. At this point the contingency between singing and rocking became intermittent when Jeanne occasionally failed to sing when the child rocked, but at this point the child took the toy back.

Kathy dropped the toy accidentally to the floor in taking it back and cried. Jeanne said, "Do you want to pick it up? I'll help you." She lifted the child off the horse onto the floor. The *child* lifted her hands. The child continued to rock without any singing from Jeanne. She dropped her doll again and the same episode as above was repeated. At this point, Jeanne probed with an even more vigorous rock of the horse. The probe caused the child to rock even more vigorously. Then Miss Simons moved the toy to the couch, which produced a slight interruption in the rocking as the child watched. The child glanced at the doll, and then all her attention was elsewhere. Jeanne picked up the doll and tapped it. The child looked at Jeanne, vocalized, and then rocked in rhythm to the tapping. Then Jeanne returned the doll to the child.

Notice again the successive approximation of a final condition in which the child is without the toy. Note the very subtle supporting stimuli, the careful programming of very brief periods, and the slow introduction of intermittent reinforcement.

> When she took the toy away, the child cried but continued rocking. Jeanne sang, which stopped the crying, and then took the child off the horse, allowing her to get the toy from the couch where they then sat together, the child on Jeanne Simons' lap. The child tried to climb on the horse by pulling Jeanne's arm in the direction of the horse. Jeanne Simons picked her up but walked away still smiling. Then, there came a period in which the child was on and off the lap. Here again probably was a very deliberate switch in contingencies in which Jeanne, having developed a repertoire of engaging Kathy via the behaviors on the rocking horse, shifted to an entirely new set of behaviors and did not reinforce any attempt to go back to the horse. Jeanne Simons moved to a small chair in a new location and took the horse along with her. Kathy stood next to the horse, played with it, and gestured to be picked up; but Jeanne placed her on the couch before putting her on the horse.

As we are discovering new ways of describing therapeutic procedures by a functional analysis of current procedures, we are finding methods of training the staff to apply them. Where the staff was already using effective procedures, our functional analysis allowed them to continue more rationally and to meet the child's behavior more frequently. Our plan at Linwood is to progressively activate the milieu--by interpersonal procedures, general rules of the institution, and devices such as teaching machines. In effect, the children teach us how to teach them; the changes in *their* behavior differentially reinforce *our* behavior. If we can describe the behavior of the children objectively and with a fine-grain correspondence between our procedures and the child's performances, small changes in the child's behavior will successively approximate effective therapeutic behavior.

In summary, I have emphasized the minimal nature of the autistic child's repertoire, the preponderance of behaviors which occur because of their aversive effect on those around the child, the general kinds of interactions with the environment that weaken a child's behavior, and some kinds of circumstances in the normal ecology that lead to extinction and other contingencies which weaken behavior. If the *deficient environments* have led to the autistic child's difficulties, then we may one day soon, from research efforts on the environmental control of operant behavior, discover procedures for reshaping his behavior.

References

Baer, D. and Sherman, J. Reinforcement control of generalized imitation in young children. Journal of Experimental Child Psychology, in press.

Bender, L. Childhood schizophrenia. Psychiatric Quarterly, 1953, 27, 663-681.

Bijou, S. Therapeutic techniques with children. In Pennington and Berg (eds.) An introduction to clinical psychology, revised edition. New York: Ronald, 1954.

Davison, G. The training of undergraduates as social reinforcers for autistic children. In Ullman and Krasner (eds.) Case studies in behavior modification. New York: Holt, Rinehart, & Winston, 1964.

Ferster, C. The repertoire of the autistic child in relation to principles of reinforcement. In Gottschalk and Auerback (eds.) Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1965. In press.

Ferster, C. Positive reinforcement and behavioral deficits in autistic children. Child Development, 1961, 32, 437-456.

Ferster, C. and DeMeyer, M. A method for the experimental analysis of the behavior of autistic children. American Journal of Orthopsychiatry, 1962, 32, 89-98.

Ferster, C. and DeMeyer, M. The development of performances in autistic children in an automatically controlled environment. Journal of Chronic Diseases, 1961, 13(4), 312-345.

Kanner, L. Early infantile autism. Journal of Pediatrics, 1944, 25, 211.

Kanner, L. Autistic disturbances of affective contact. Nervous Child, 1943, 2, 217-250.

Lovaas, I. A recording method and observations of behaviors of normal and autistic children in free play settings. Journal of Experimental Child Psychology, in press.

Lovaas, I., Freitag, G., Gold, V., and Kassorla, I. Experimental studies in childhood schizophrenia: Analysis of self destructive behavior. Journal of Experimental Child Psychology, 1965, 2, 67-84.

Lovaas, I., Freitag, G., Kinder, M., Rubenstein, B., Schaeffer, B., and Simmons, J. Experimental studies in childhood schizophrenia: Establishment of social reinforcers. Paper read at Western Psychological Association, Portland, 1964.

Lovaas, I., Schaeffer, B., and Simmons, J. Experimental studies in childhood schizophrenia: Building social behaviors in autistic children using electric shock. Journal of Experimental Studies of Personalities, in press.

Sherman, J. Use of reinforcement and imitation to reinstate verbal behavior in mute psychotics. Journal of Abnormal Psychology, in press.

Wolf, M., Risley, T., and Mees, H. Application of operant conditioning procedures to the behavior problems of an autistic child. Behaviour Research Therapies, 1964, 1, 305-312.

CHAPTER XI WHAT BEHAVIORAL ENGINEERING IS[1]

L. E. Homme

The phrase "behavioral engineering" has been around a long time and all of us understand it. The only trouble is--we haven't told anyone, including ourselves, what it means. It probably is true that everyone would agree that behavioral engineering is "arranging the environment so that one gets the behavior one wants,"[2] or, more generally that behavioral engineering is the application of laws of behavior to practical problems. Nevertheless, the topic of what behavioral engineering is could stand some elaboration. This becomes most apparent when one sets out to shape up a behavioral engineer. Then it quickly develops that behavioral engineering is a blend of two technologies: the technology of contingency management and the technology of stimulus control. Needless to say, no special degrees or certificates are required to put the technologies to use. The laws of nature do not care about such matters; they go on working in any case.

In analyzing a behavioral engineering task, the engineer must determine: (a) exactly what behavior he wants to occur, (b) what stimuli are to control it, and (c) what reinforcers are available.

The technology of contingency management deals with managing reinforcers (c), and the technology of stimulus control with arranging stimuli so that

Originally appeared as a reprint in The Psychological Record, 1968 18, 425-434, with Polo C'de Baca, Lon Cottingham, and Angela Homme.

they come to control (b). Society at large generally decides on what behaviors should occur (a), but the time may have arrived when behavioral engineering itself may have something to say. The relationships between the technology of contingency management and the technology of stimulus control and behavioral engineering may be summarized by relating them to the familiar three-term contingency:

Discriminative Stimulus → *Performance* → *Reinforcing Event*

Technology of Stimulus Control　　　*Technology of Contingency Management*

Behavioral Engineering

Figure 1. Relationships between the three-term contingency and behavioral engineering.

Clearly it is possible to be a contingency manager and understand little about stimulus control, but the reverse certainly is not true. In order to bring a behavior under stimulus control, contingencies have to be properly managed.

The Technology of Contingency Management

Contingency management is the management of what events are contingent upon what behavior. It is clear that contingency management is merely the taking seriously (literally) that great law of life: When reinforcing events are contingent upon a given behavior, the behavior will increase in strength; when they are not, the behavior will decrease in strength. The power and generality of contingency management is by now becoming clearer and clearer (Ferster & Perrott, 1968; Homme, 1967; Krasner & Ullman, 1965; Ullman & Krasner, 1966).

The great law of life is simple, but this fact ought not be allowed to obscure another fact: The law is powerful.

Also, the simplicity of the overriding consideration of contingency management (that behavior depends on its consequences) should not lead one to believe it is always simple to teach; some trainees grasp the principle and are able to put it to work instantly; others, in our experience, never.

The critical ingredient appears to be an emotional commitment or willingness to *pay off* for desirable behavior. With this commitment, things proceed apace; without it, nothing good happens.

One can make a pretty good case that, basically, there are only two things that a good contingency manager has to know and do: (a) to reinforce the behavior he wants, and (b) to recognize and reinforce *approximations* to this behavior.

The Relationship Between Operant Conditioning and Contingency Management

Everybody strengthens, weakens, or suppresses operants, and everybody, in this sense, is an operant conditioner. And, since everybody at some time or other, arranges consequences for behavior, everyone is a contingency manager. Assuming that we all agree that these usages are too broad to be useful, let us get on to narrowing them down.

Let us reserve the phrase "operant conditioning" to designate the basic laboratory science from which behavioral engineering is derived. This means, then, that an operant conditioner is an operant conditioner only as long as he remains in his laboratory. When he leaves the laboratory and systematically applies the *principles* of behavior to problems like getting his offspring to pass a progress check on his homework before he gets points which can be used to buy reinforcers, the operant conditioner becomes a behavioral engineer. We are calling him a behavioral engineer rather than a contingency manager, because we are assuming he knows the rules underlying the technology of stimulus control as well as those underlying contingency management. From this standpoint then, the nurses whom Ayllon and Michael (1959) called behavioral engineers should more properly be called contingency managers; Ayllon and Michael were the behavioral engineers. Similarly, in the Gelfands and Dobson study (1967) of who reinforces what in a mental hospital, the patients, nurses, and nurses' assistants were contingency managers (albeit poor ones), not behavioral engineers. In the same view, when Ogden Lindsley (1960) did his pioneering work on psychotics pulling plungers in an experimental enclosure, he was an operant conditioner, but now that he is teaching teachers and parents to teach (Lindsley, 1966), he is engaged in behavioral engineering. Both Lindsley and Homme were once operant conditioners, but they quit.

It may be worthwhile to mention some differences between the operant conditioning laboratory model and contingency management which employs the *principles* of operant conditioning but not the *intact model* of the operant conditioning laboratory. To elaborate: In the animal operant conditioning laboratory, one usually is working with a deprived organism; the deprivation operations serve to insure that one class of events will be a reinforcer--the presentation of the event of which the beast was deprived. The reinforcing event, e.g. presentation of food, is thus stable and highly effective as a reinforcer both within and between experimental sessions. With non-deprived humans, the model is often unsatisfactory; reinforcers often wear out and lose their reinforcing properties quickly. Because of this, we have found Premack's differential probability hypothesis, which states that *any* response can reinforce *any other* response of lower probability, of immense usefulness (Homme, 1966; Premack, 1959). We find that the differential probability hypothesis combined with the use of a reinforcing event menu, on which many reinforcers are listed or pictured (Addison & Homme, 1966), highly useful in generating tens or hundreds of reinforcers where we had one or two before. By having S select his reinforcer from a menu, one determines precisely what is the most effective reinforcer here and now for this S.

The Technology of Stimulus Control

A behavioral engineer's definition of stimulus control is a simple one: Stimulus control exists to the extent that the presence or absence of a stimulus controls the probability of a response. It is interesting to compare this with the basic researcher's definition: "Stimulus control refers to the extent to which the value of an antecedent stimulus determines the probability of occurrence of a conditioned response" (Terrace, 1966, p. 271). The key difference in these two definitions lies in the words, *extent* and *values*. It is clear that the behavioral engineer is really interested in approximations to only two probability values, 1.0 and 0.0; the basic researcher, on the other hand, has to worry about values between these two values--the slope and shape of generalization gradient.

Faulty Stimulus Control

"He won't mind!" "He's stubborn!" "He knows how to do it, but he doesn't feel like it." "I meant to do it, but I forgot." In all these cases, one is speaking of faulty stimulus control. As a matter of fact, one could make an excellent case for the fact that *most* behavioral engineering problems are problems of faulty stimulus control; that is, the S has the response in his repertoire, all right, but it is not made when the stimulus is presented.

Correcting Faulty Stimulus Control

The fundamental rule for correcting faulty stimulus control is the same as that for establishing stimulus control in the first place. Get the behavior (or some approximation to it) made while the S is attending to the stimulus which is to control it. This means, among other things, that the engineer must be able to reinstate the control stimulus whenever he wants to.

Using the Behavior of the Subject to Reinstate the Discriminative Stimulus

Over 30 years ago, Guthrie (1935, p. 18) published the following story about stimulus control. It may be time to take it seriously.

> The mother of a ten-year-old girl complained to a psychologist that for two years her daughter had annoyed her by a habit of tossing coat and hat on the floor as she entered the house. On a hundred occasions the mother had insisted that the girl pick up the clothing and hang it in its place. These wild ways were changed only after the mother, on advice, began to insist not that the girl pick up the fallen garments from the floor but that she put them on, return to the street, and reenter the house, this time removing the coat and hanging it properly.

With the usual perspicacity of mommies, the lady for two years had been getting the response out in the presence of the wrong stimuli. The stimuli which were supposed to control the response were those prevailing immediately after the child entered the house; instead, the response was repeatedly evoked in the presence of some stimulus such as, "Please pick up your coat."

To show that the laws of behavior don't change every 30 years or so, it is worth examining one more example. Evans relates that, during a visit, his four-year-old niece was exhibiting a minor, but irritating, behavior disorder.[3] The child, who was nicely toilet trained, never seemed able to remember to flush the toilet at the appropriate time. Instead, she would leave the bathroom and have to be ordered back to do that job.

This problem, it turned out, had persisted for at least a couple of years until finally Uncle Jim intervened. Instead of ordering the child back to flush the toilet, he instructed her to make believe that she was just finishing going to the toilet; she was to make believe to the extent that she was to take down her pants, climb up on the toilet, imagine she was just finishing, get off the toilet, pull up her pants, then flush the toilet and rejoin the adults. This exercise worked; the child gives every sign that she will grow up to be a happily married lady who flushes the toilet a lot.

The Effectiveness of Guthrie's Technique

There are at least two considerations which account for the effectiveness of Guthrie's technique of getting the S to go back and run off some of the behavior chain which leads up to the point of difficulty. The first is that the effect of a stimulus persists for some time after the stimulus is terminated. The second is that an important portion of the stimuli which are going to control is response-produced.

Both of these factors can be seen at work in the case of Guthrie's example. The effects of the "outdoor stimuli" obviously do not terminate the instant the child crosses the threshold. If it is a brisk day, for example, she may still be feeling the effects of the cold, she may be panting from running, and the little tyke may even be saying sentences to herself such as, "Christ, it's cold today."

It is worth noticing in the same example that all of the stimulus changes under discussion were produced by the behavior of the organism: The visual and other changes brought about when the S opened the door; the changes in stimulation caused by the removal of the coat and dropping it on the floor; the other changes in visual stimuli caused by the child's moving her head or her eyes. Note that a description of proprioceptive stimuli, as important as they may be, was not required in this discussion of response-produced stimuli. This is worth noting because sometimes proprioceptive stimuli and response-produced stimuli tend to get equated when, of course, they should not be.

In summary, then, the technique of requiring the S to run off the part of the behavioral chain just preceding the response which is giving difficulty is as successful as it is, because it is a way of insuring that the stimulus which is to control the response is, in fact, reinstated. It is important to notice that this is true whether or not the behavioral engineer can specify all of the portions of the stimulus complex which are likely to be present when the response is to be executed; running off the immediately preceding chain members automatically guarantees the presence of the control stimuli.

An Example of Behavioral Engineering: Teaching Phonic Reading

There are children who cannot read, but whose parents and/or teachers insist that they "know phonics." It may be true that they know phonics in the sense that, shown any letter, they can give its sound. Thus, to the printed word stimulus "cat" the student can respond, "cuh-aa-tuh." It is from here on that trouble arises. To the self-generated stimulus, "cuh-aa-tuh," he may respond "supersonic transport," or whatever else happens to be at high strength at the moment. Analysis of phonic reading, then, yields the kinds of chains shown in Figure 2.

Visual Stimulus → *Vocal Response* → *Auditory Stimulus* → *Vocal Response*

Cat → *"cuh-aa-tuh"* → *"cat"*

Figure 2. Sequence of events which maintains behavior and insures appropriate stimulus control.

In the S just described, it is obvious that he simply had not learned the last member of the chain--the translation of "cuh-aa-tuh" to "cat." In the laboratory, the last member of the chain is the first one taught (Homme & Klaus, 1957); so we adopted the same strategy in teaching reading. The first thing taught to the child is the "sound-the-word game." This simply consists in making reinforcing events contingent upon the child's translating phoneticized words. For example, the contingency manager may say to the child, "Tell me what I'm saying, 'chuh-air'." The child who knows the game will answer, "Chair." The contingency manager then says, "Good. Choose what you'd like to play with from the menu." It has been our experience that very shortly the child will begin to make up phoneticized words with which to puzzle the contingency manager.

Once the end of the chain (the translation skill) has been established, it is a simple matter to move up the chain, step by step.

With these kinds of procedures, it has been possible to keep preschool children responding eight hours a day and to teach phonic reading in a matter of days, rather than semesters or years.

Behavioral Engineering Research

The behavioral engineering field is about at the same stage of development as Goddard's rocket program was in 1935 (Lehman, 1963). At this time, those who know about such matters say that not only had the basic research been done, but plenty of technology was available to put an object into space. (Goddard had already blasted a rocket up a thousand feet.) What was required, and what took our society 20 or so years to realize, was a real effort at implementation.

Just as space scientists and engineers didn't have to wait until the "gravitational process" was better understood before the law of gravity could be exploited, so behavioral engineering need not wait until the "learning process" is better understood before we can get on with some important jobs. For openers, we can develop a technology for routinely producing superior human beings. We may have sufficient technology--here and now--(Homme, 1967) to be able to guarantee that, given a physiologically normal human being, and given control of his reinforcement contingencies and stimulus conditions, we can shape him into a superior organism. He will be superior not only intellectually, he will be superior emotionally. He will be happier than most, have a better self-concept--he will have a better repertoire in all the ways we can think of to make it better.

Another way of saying this is that we have the technology for installing any behavior we want. The problem now is, what behaviors do we want installed. We submit that some leading condidates are a preschool academic repertoire, a favorable self-concept, love, and joy.

The Intellectual Repertoire of the Preschool Child

No one knows how much a preschool child can learn. This is so despite the thousancs of developmental studies on how much children *usually* learn by a given age. There is no research which has, in a systematic manner, explored the limits of a child's capacity to learn when contingencies were intelligently managed, and stimulus control criteria were met.

Since behavioral engineering views reading as discriminative responding, the ridiculous problem of "When is the child ready to read?" never arises. From a behavioral engineering standpoint, a child is reading objects as soon as he can discriminatively respond to them (e.g., name them). Once a child has begun to read, no one knows, really, how fast his education may proceed. With the opportunity to read coupied with unsystematic contingency management, one scholar who learned to read at age three read at the fifth-grade level by the time she was five.[4] (Her brother, the control group, learned to read exactly when educational researchers said he should--at 6.5 years of age.) Once having learned to read, even at a primitive level, the child can begin going through existing programed instructional materials--say, in arithmetic. With contingency management, the excuse that existing programs are boring, dull, or no good, is not valid. One does not hear the rat in the Skinnerbox or the human at the slot machine complain about the dullness of his task. He eagerly does his job, although he may have done the same thing hundreds of times before. And programmed instruction, even though bad, does offer more variety than slot machines; it is the payoffs which have to be arranged.

And even if stimulus control in the program is faulty--as is often the case--this can be corrected by means of progress checks. Progress checks are short tests on material just covered in the program--criterion items which should have been, but were not, built into the program. The sequence of events shown in Figure 3 thus results.

N Frames of Program ⟶ *Progress Check* ⟶ *Reinforcer*

Figure 3. Stimulus-response analysis of a beginning phonic reading sequence.

With a schedule of events of this sort, no one knows how fast a child's intellectual development might proceed. We do have enough preliminary data to know that there is no danger in "pushing" a child too fast with such a system. Three-year-olds have been kept working eight hours a day with no signs of the "I don't want to do this any more" response. On the contrary, having found a place where someone was willing to pay off for desired behavior rather than punish for unwanted behavior, they wanted *more* of the same at the end of the day--to the despair of the worn-out contingency managers.

Joy

Psychologists have been assiduous in studying the unpleasant aspects of life. Pain, depression and anxiety have received considerable attention, but the same cannot be said for joy and happiness. No "Manifest Joy Scale" exists. The point is, the serious study of joy is overdue.

Skinner (1953, p. 127) has observed that the emotion called joy involves the whole repertoire of the organism. ". . . our [joyful] subject speaks to everyone, reacts in an exaggerated fashion, walks faster and seemingly more lightly, and so on. This is particularly obvious in the behavior of young children--for example, on the eve of a holiday or festival."

It may be that the overwhelming nature of this fact--that joy involves the whole repertoire--is what has impeded the study of joy. Behavioral engineering certainly has no tools for dealing directly with a whole repertoire at once. But if we take seriously what Keller and Schoenfeld (1950) and Skinner (1953) have said about what joy is, we can reduce the joy problem to manageable size. They say that joy is the anticipation of reinforcing events. This simple, but profound, observation is clearly borne out by the amount of laughter and smiling which occurs when children's contingencies are planfully managed. Reinforcements are frequent, and there is little doubt that the child quickly learns to anticipate reinforcing events and to plan what he will next select from a reinforcing event menu.

In order for the behavioral engineer to strengthen the joy response, then, he calls for it immediately preceding a reinforcing event. E.g., to a child, "Tell me something good that's going to happen to you; then go to the menu and choose what you'd like to do."

In less formal situations the same rules hold: Make the S pay for reinforcing events by verbalizing the anticipation of other reinforcing events. "Tell me something good that you're going to do, and then we'll go to the store."

Of course, since everyone is an organism, one can try this out on oneself. One can increase the frequency with which one coverants joy events (Homme, 1967) and observe the effect. (The word "coverants" is a contraction of covert operant, and is pronounced "kuhverant"; when used as a verb, it means "think about," "imagine.") A convenient property of self-management research is that one always has a S close at hand.

Love

Joy is difficult to define, but perhaps even more difficult is love. Many famous writers, for example, have written a great many words to persuade the reader that the love *he* is talking about is indescribable. That may be so, but the behavioral engineer, assuming that love in the repertoire is desirable, must get on with the job of trying to install it.

If one examines the behavior of a human who is said to exhibit love, one can quickly detect one very public behavior. This is verbal behavior with the verb "love" in it. Assuming that this is a genuine signal of love, or an approximation to it, it is a straightforward matter to install this kind of verbal behavior. E.g., the following class of verbal behaviors can be easily and quickly installed in a S's repertoire.

"Why do you love Becky (the contingency manager)"?

"I love Becky because . . ."

If one also accepts the proposition that love is manifested by a tendency to reinforce, then another approximation would be to evoke and strengthen reinforcing verbal behavior in the S. "Go whisper to Becky that you love her because . . ." Observations suggest that verbal behavior such as this from a preschool S can turn an adult on. This state of affairs makes it almost unnecessary to instruct the contingency manager to reinforce. We have the impression that she couldn't refrain if she tried.

A Favorable Self-concept

This also is a complex phenomenon, we are told. How can you teach someone to think favorably of himself? Put this way, the question may indeed seem formidable. However, analysis leads to the same conclusion as that to which Ellis (1958) has come. When we speak of a self-concept, we may simply be talking about the aggregate of sentences the S says *to* himself (and others) *about* himself. Viewed in this light, it becomes a simple matter to install a favorable self-concept: One simply strengthens a class of verbal behavior. E.g., after a good performance, "That was a fine job. Go whisper to your teacher, 'I am a fast learner'." At this point, the teacher reinforces verbally and by making the menu available. It has been our experience that after a very few trials of this sort, the "I am a fast learner" conception of himself competes quite successfully with the child's "I sure am dumb" concept.

A similar technology can be brought to bear in psychotherapy. For example, the "poor, helpless me" self-concept is replaced by the self-mastery statement, "I'm in charge of my own behavior," by having S program himself so that this coverant occurs with high frequency.

There are many, many fascinating engineering research areas which have not been touched upon, of course. For example, Bandura's (1964) modeling concepts, if taken seriously, suggest that, before a great child can become a great adult, he has to know a great adult. We may have to face up to the fact that the Fred Skinners of the world are not spread thin enough.

Footnotes

1. We thank Rebecca Robinson and Gary Keys for their creative and energetic technical assistance and Thomas Topich, departmental secretary and librarian, for coming to work almost every day.

2. R. E. Ulrich, personal communication, 1967.

3. J. L. Evans, personal communication, 1965.

4. J. L. Evans, personal communication, 1967.

References

Addison, R. M., & Homme L. The reinforcing event (RE) menu. NSPI Journal, 1966, 5 (1), 8-9.

Ayllon, T., & Michael, J. The psychiatric nurse as a behavioral engineer. Journal of the Experimental Analysis of Behavior, 1959, 2, 323-334

Bandura, A., & Kupers, C. J. Transmission of patterns of self-reinforcement through modeling. Journal of Abnormal Social Psychology, 1964, 69 (1), 1-9.

Ellis, A. Rational psychotherapy. Journal of General Psychology, 1958, 59, 35-49.

Ferster, C. B., & Perrott, M. C. Behavior principles. New York: Appleton-Century-Crofts, 1968.

Gelfand, D. M., Gelfand, S., & Dobson, W. R. Unprogrammed reinforcement of patients' behavior in a mental hospital. Behaviour Research Therapies, 1967, 5 (3), 201-207.

Guthrie, E. R. The psychology of learning. New York: Harper, 1935.

Homme, L. Contingency management. Newsletter, Section on Clinical Child Psychology, Division of Clinical Psychology, American Psychological Association, 1966, 5 (4).

Homme, L. A behavior technology exists--here and now. A version of this paper was given at the Aerospace Education Foundation's "Education for the 1960s" Seminar, Washington, D.C., 1967.

Homme, L., & Klaus, D. J. Laboratory studies in the analysis of behavior. (3rd edition) Albuquerque: Westinghouse Learning Corporation, 1967.

Keller, F. S., & Schoenfeld, W. N. The principles of psychology. New York: Appleton-Century-Crofts, 1950.

Krasner, L., & Ullman, L. P. Research in behavior modification. New York: Holt, Rinehart and Winston, 1965.

Lehman, M. This high man: The life of Robert H. Goddard. New York: Farrar, Straus, 1963.

Lindsley, O. R. Characteristics of the behavior of chronic psychotics as revealed by free-operant conditioning methods. In Diseases of the Nervous System, (Monograph Supplement), 1960, 21 (2).

Lindsley, O. R. An experiment with parents handling behavior at home. Johnstone Bulletin, 1966, 9, 27-36.

Premack, D. Toward empirical behavior laws: I. Positive reinforcement. Psychological Review, 1959, 66, 219-233.

Skinner, B. F. Science and human behavior. New York: Macmillan, 1953.

Terrace, H. S. Stimulus control. In W. K. Honig (Ed.), Operant behavior Areas of research and application. New York: Appleton-Century-Crofts, 1966, pp. 271-344.

Ullman, L. P., & Krasner, L. (Eds.) Case studies in behavior modification. New York: Holt, Rinehart and Winston, 1966.

CHAPTER XII

A CONTINGENCY MANAGEMENT CLASSROOM: BASIS FOR SYSTEMATIC REPLICATION

T. C. Lovitt

Systematic replication has generally been defined as a series of experiments--one research effort leading to another, with each succeeding experiment varying to some extent from the previous one. Perhaps some new independent variable is manipulated or the topography of the response is changed. In these aspects, it differs from direct replication, where an experiment is simply repeated, using the same experimental procedures with either the same or a different subject.

Direct replication might be likened to the repeated rereading of one chapter of a book. Presumably the reader, with each reading, steadily increases his knowledge of the content of that particular chapter. Systematic replication is analogous to reading an entire book. In this case the first chapter gives meaning to the second, the second to the third, and so on, until, finally, when all the chapters are covered, a message comprised of many constituent elements emerges.

For a number of reasons, this latter form of experimentation has been used almost solely by the experimental psychologist working in a carefully controlled laboratory, and usually with subhuman subjects. Ordinarily, the only determinants of his experimental pursuits are time, space, and budget. He has in most cases solved the problems that preclude a series of systematic investigations: response definition, stable environment, a means for recording behavior, and an effective reinforcer. Responses are generally

With Karen A. Curtiss, Co-author, in whose classroom the studies were conducted.

When systematic replication, or experimentation of any type, is attempted in classrooms, however, certain problems arise. Reading, writing, and aritnmetic responses can be vague and illusive, certainly less definable than an electric pulse. Environmental stability is even more difficult to obtain in that classes generally are comprised of many organisms exhibiting numerous and varied behaviors, many of which are controlled by unknown or weak discriminative stimuli. The classroom researcher seeking reinforcers for his pupils must be more resourceful than many laboratory scientists, for mash and pellets of grain are not generally considered accelerating to well-fed children.

Although the classroom cannot, and perhaps should not, duplicate the rigid, inflexible atmosphere of the experimental chamber, a fairly reliable situation where significant research is carried out can be provided.

Response Definition

Perhaps the first step toward setting up a classroom in which to conduct systematic research would be to use materials that lend themselves to quantification. The researcher must be able to count the mathematics, reading, spelling, or writing responses emitted by the pupil. Programed sequenced of many types are desirable in that the materials are divided into periodic elements of frames. Furthermore, the steps in many programed materials are in relatively equal units. In some subject areas, where programed materials are unsatisfactory or do not exist, the classroom manager himself must subdivide the curriculum so that some equality from one element to the other is assured.

Stable Environment

In order to secure meaningful response data, the classroom manager must set up the classroom in such a way that the physical environment is relatively the same from day to day. He must arrange chairs, study carrels, and instructional centers to accommodate the presentation of materials and the confirmation of responses, changing the arrangement only when deemed functionally necessary. Further stability in a classroom can be achieved by consistency in the scheduling of the various curricular activities, such as presenting instructions and materials, confirming answers, and requesting information.

Recording Responses

Some method must be established to procure response data. This is largely solved by using programed materials that require the pupil to make a pencil response. Thus, the classroom manager has available a permanent response record. In addition, data recording sheets could be provided that would enable either the teacher or the pupil to translate the data from the material itself, thereby furnishing the researcher with a convenient survey of response data over a period of time.

Establishing Reinforcers

A variety of activities and exercises in the classroom can be used on a contingent basis. Such activities, ranging from recess to library books, are limited only by the imagination and engineering capabilities of the classroom manager. Several classroom managers have found that the most efficient and precise way to use such reinforcers is to require the pupil to purchase an activity with a certain number of marks, points, or tokens. The pupil, then, is granted so many tokens contingent upon a specified number of academic responses. With these tokens he can periodically purchase all or a portion of an activity.

The following four experiments demonstrate systematic replication. They were conducted in a contingency management classroom in which programed materials were used almost exclusively. Academic responses were recorded on data sheets, and point-per-answer contingencies were employed--points being redeemable for leisure time activities.

The first study compared the effects of teach- and pupil-specified contingency ratios on academic rates. Since the choice of contingencies ordinarily falls to the teacher, this experiment is an assessment of one phase of self-management.

The second investigation used a manipulation of the ratios at which contingencies were delivered. During two of the experiments, a student received a uniform contingency of one minute of free time for every twenty correct math responses. For two other experiments, reinforcement was given in ratios that increased as the subject worked at a higher rate.

An assessment of teacher training is presented in the third study. A student was trained by his teacher to tutor another member of his class. Both the tutor alone and the student and tutor together were evaluated on the basis of the student's progress.

The purpose of the last study was to assess the effects of verbalizing a math problem prior to making a written response. Results showed that as a result of the verbalization, correct rate increased and error rate decreased.

Study I: Academic Response Rate as a Function of Teacher- and Self-Imposed Contingencies

The management of one's own behavior has often been expressed as one of the prime objectives of our educational system. The problem, however, in programing toward such an end has been one of definition--specifying those variables that constitute self-management. Until the skills or traits that lead to a self-managing individual are clearly detailed and explicitly defined, the objective of self-management may never be realized. If, however, certain sub-properties of self-management can be determined, sequentially ordered, and systematically presented, the probability of realizing this objective is greatly increased.

An individual who can control or manage his own behavior may be a person who has the ability to assess his own competencies, set his own behavioral objectives, and specify a contingency system whereby he might obtain these objectives. Translated to a school situation, this would be an individual who knew his academic capabilities in terms of skill levels and rate of performance, could arrange a series of activities or steps to achieve a variety of self-imposed objectives, and could grant himself reinforcement on a prearranged schedule to accomplish certain behavioral sequences.

This last behavior, self-specification of contingencies, has been a neglected area of investigation. Traditionally, it is the teacher who arranges the contingencies in the classroom, saying to the pupil, "If you do these arithmetic problems, you will receive a gold star," or "If you correctly answer these questions you will be allowed to go out for recess." Infrequently is the student himself allowed to arrange his own educational environment.

The purpose of this study was to analyze functionally the effects of self-imposed versus teacher-imposed contingencies. In this study the specification of contingencies was the independent variable, while academic response rate was selected as the dependent variable.

Subject and Conditions

The subject in this study was a 12-year-old member of a class for children with behavioral disorders at the Experimental Education Unit, University of Washington. The subject had been a member of this class, whose management system was based on the Premack design of high and low probability behaviors (Haring & Kunzelmann, 1966; Haring & Lovitt, 1967), for two academic years. Contingent upon academic responses (low probability behaviors), the subjects in this class were given points in each of the academic areas. These points were later converted to minutes of time in the high interest (high probability behavior) room. The ratios of points-per-answer were individually specified, not only among class members but also among subject matter areas for any one student.

Procedures

This investigation consisted of three separate experiments--two that manipulated the contingency manager and one that manipulated magnitude of reinforcement. During Experiment I baseline data relevant to the subject's academic response rate were obtained for nine days. Each day a response rate figure was calculated that represented the subject's performance in all of his scheduled subject matter areas. Throughout this period no attempt was made to explain to the subject the response-per-point ratio in each academic area.

Following this baseline period, Stage I of the study was instituted. It was the intent at this time to instruct the subject about the relationship between correct answers and contingent points. Each day in this 12-day stage the teacher verbally explained the contingencies and placed a written copy of them on the subject's desk. Table I outlines the contingency system in effect throughout Experiment I. The contract was composed of 9 agreements, each of which had a response-per-point ratio. For example, the subject was granted 2 points for each page read (Sullivan Associates Program, 1963) and I point for 10 correct mathematics problems (Singer Mathematics Program, 1965). As the subject completed each academic assign-

Table I

Contingency Specifications

During Experiment I

Subject Area	Contingency Manager: Teacher-Specified	Contingency Manager: Child Specified
Math	10 problems: I minute free time	10 problems: I minute free time
Supplementary Math	10 problems: I	10 problems: I
Reading		
(No errors)	I page: 2	I page: 2
(Errors)	I page: I	I page: I
Spelling	I page: I	*10 words: 2
Writing	20 letters: I	*10 letters: 2
Library Reading	I story: 3	*I story: 10
Cyclo-Teacher		
Multiplication	I side: I	I side: I
Spelling	I side: 2	I side: 2

*Indicates where the child-imposed requirements differ from those imposed by the teacher ment, he was shown how many responses had been made and was asked to calculate the corresponding points he had earned.

In Stage 2, which extended for 22 days, the copy of the response-point requirements was removed from the subject's desk. He was now asked to specify verbally his own payment in each of the nine areas and to record his decisions, which were then attached to his desk. As in Stage 1, when each assignment was completed the subject was asked to calculate the points he had earned. Finally, in Stage 3, which lasted for seven days, the teacher-imposed contingencies were again in effect.

In Experiment I, as in all experiments of this study, the sequence of academic areas was basically the same each day; reading was followed by math, then spelling, English, and writing. Furthermore, the time allotted to each subject was about the same each day--two hours for reading, one hour for math, and a total of one and one-half hours for spelling, English, writing, and library reading.

Regardless of any variability in the ordering of the subject areas or the time allotted to them, however, it was always the teacher, not the subject, who arranged the day's academic program. Thus, the subject could only work on each academic activity as it was scheduled--math during the math period and reading during the reading period. He could not switch freely from one academic area to another, regardless of the contingencies that were in effect from one activity to another or who had imposed the contingencies.

After a period of four weeks, which separated the academic quarters at the University of Washington, Experiment II was begun. No baseline data were obtained during this second experiment, since the subject was now fully acquainted with the response-per-point contingencies. Other procedures were carried out as before, in order to replicate Experiment I.

Teacher contingencies were explained, written out, and attached to the subject's desk in Stages 1 and 3, while during Stage 2 the subject's contingencies were in operation. The only difference between Experiments I and II was that in the first experiment 9 specific agreements were involved, while in the latter investigation 8 were included. In both experiments, however, the subject's complete program was included in the study. The response-per-point requirements of Experiment II are presented in Table II.

Table 2

Contingency Specifications

During Experiment II

Subject Area	Contingency Manager: Teacher	Contingency Manager: Child
Math	10 problems: 1 minute free time	*10 problems: 2 minutes free time
Reading		
(No errors)	1 page: 2	*1 page: 3
(Errors)	1 page: 1	*1 page: 2
Spelling	1 page: 1	*5 words: 1
Writing	20 letters: 1	*10 letters: 2
Language Arts	10 answers: 1	*10 answers: 2
Library Book	1 story: 3	*1 story: 6
	3 questions: 1	*3 questions: 2

*Indicates where the child-imposed requirements differ from those imposed by the teacher

Following this replication study, Experiment III was conducted. Since during Experiment II the subject had altered all the teacher-imposed requirements to grant himself more points per response, it was necessary to determine whether self-contingencies had affected the academic rate increase or whether this increase was due to the increased pay off. Experiment III, therefore, consisted of three stages: (1) the teacher specified the response-per-point requirements she had placed in effect throughout Experiments I and II; (2) the teacher specified the requirements that the subject had instituted during Experiment II; and (3) the teacher again specified her original requirements. These requirements were identical to those listed in Table 2. The only difference between Experiments II and III was that in Experiment III the teacher imposed the contingency requirements throughout, whereas in Experiment II the subject set his own contingencies during Stage 2.

Results

Experiment I

During the baseline period (unspecified contingencies) the subject's median rate of response was 1.8 while his median rate during Stage 1 was 1.65. The response range during the baseline observations, which lasted for 9 days, extended from 1.1 to 2.4. Stage 2 consisted of 22 daily sessions; in this stage the subject's median rate of response was 2.5, ranging from 1.4 to 3.6. A median of 1.9 was obtained in the final stage of the experiment, which consisted of 7 daily sessions, while the subject's performance extended from 1.0 to 2.2 responses per minute. The data from this experiment are displayed in Figure 1.

Experiment II

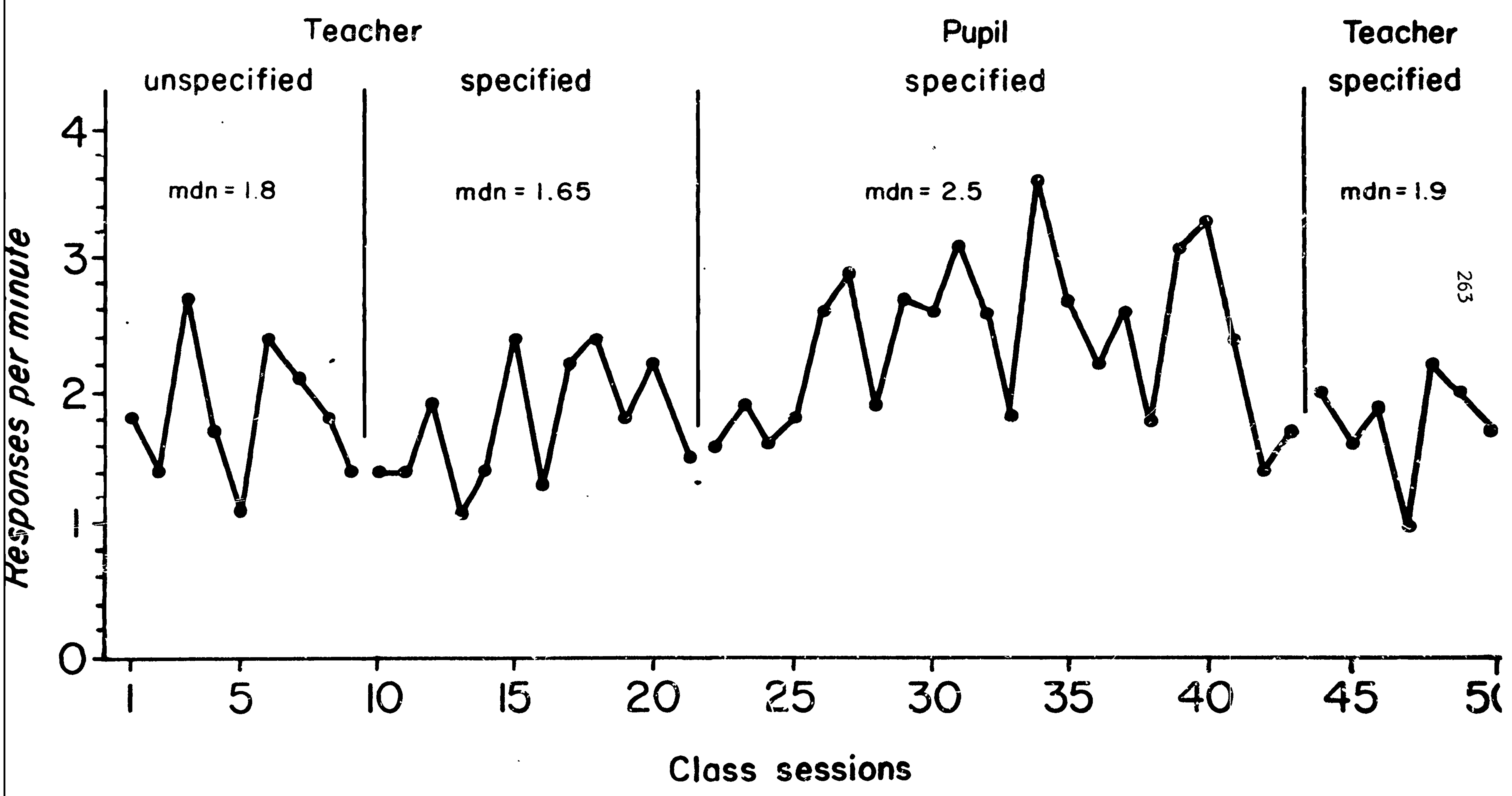
Throughout the 13 days of the first stage, the subject's rate of response ranged from .8 to 2.1 with a median response rate of 1.6. During the second stage of this experiment, which also lasted for 13 days, the subject's medain rate was 2.3 responses per minute. His range at this time was 2.0 to 2.6 responses per minute. In the 9 days of the final stage, the return to teacher contingencies, the subject's median rate of response was 1.5 and his range, 1.2 to 1.7. The results of Experiment II are shown in Figure 2.

Experiment III

The data from the final experiment are presented in Figure 3. The

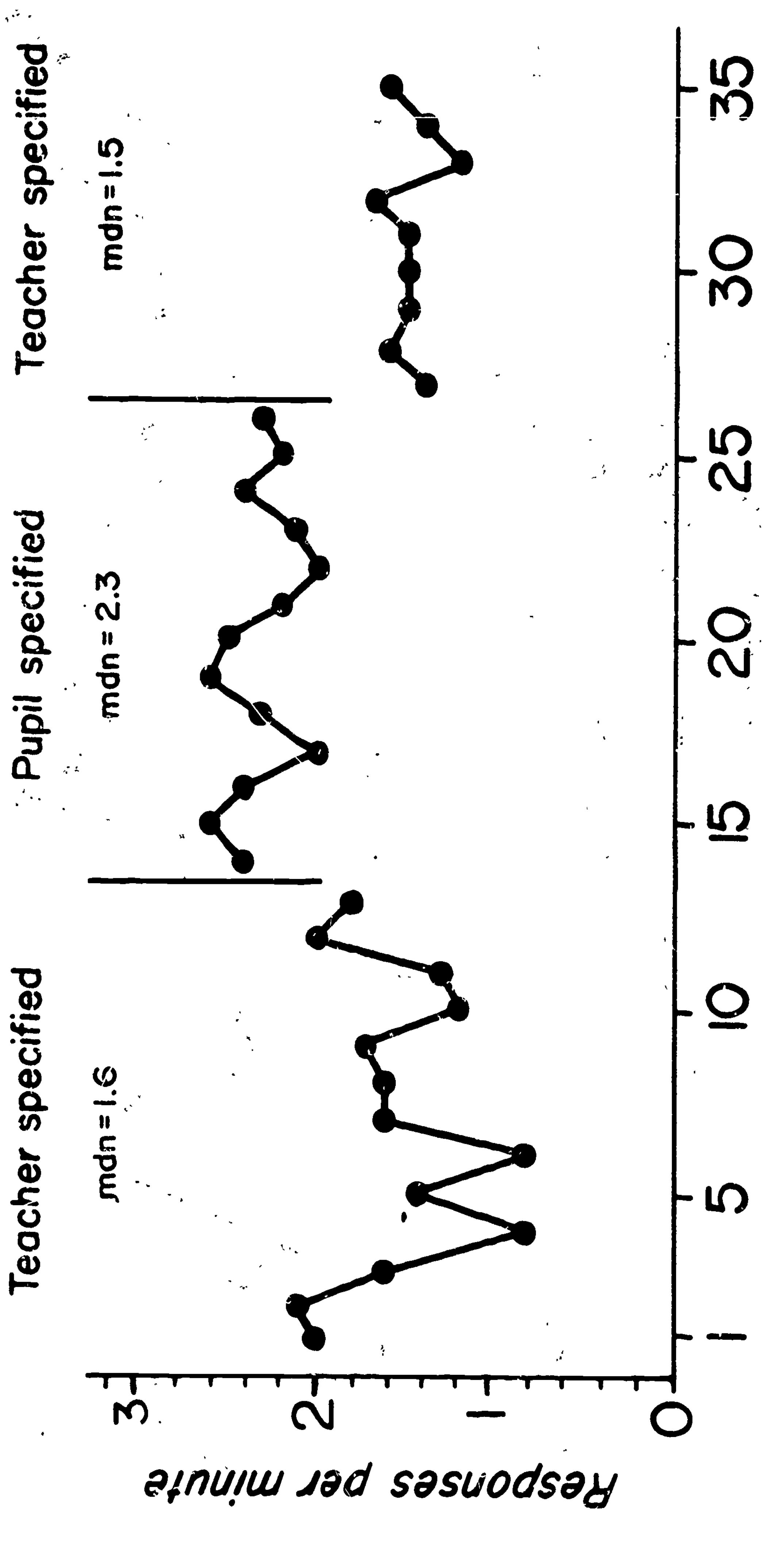
Experiment I

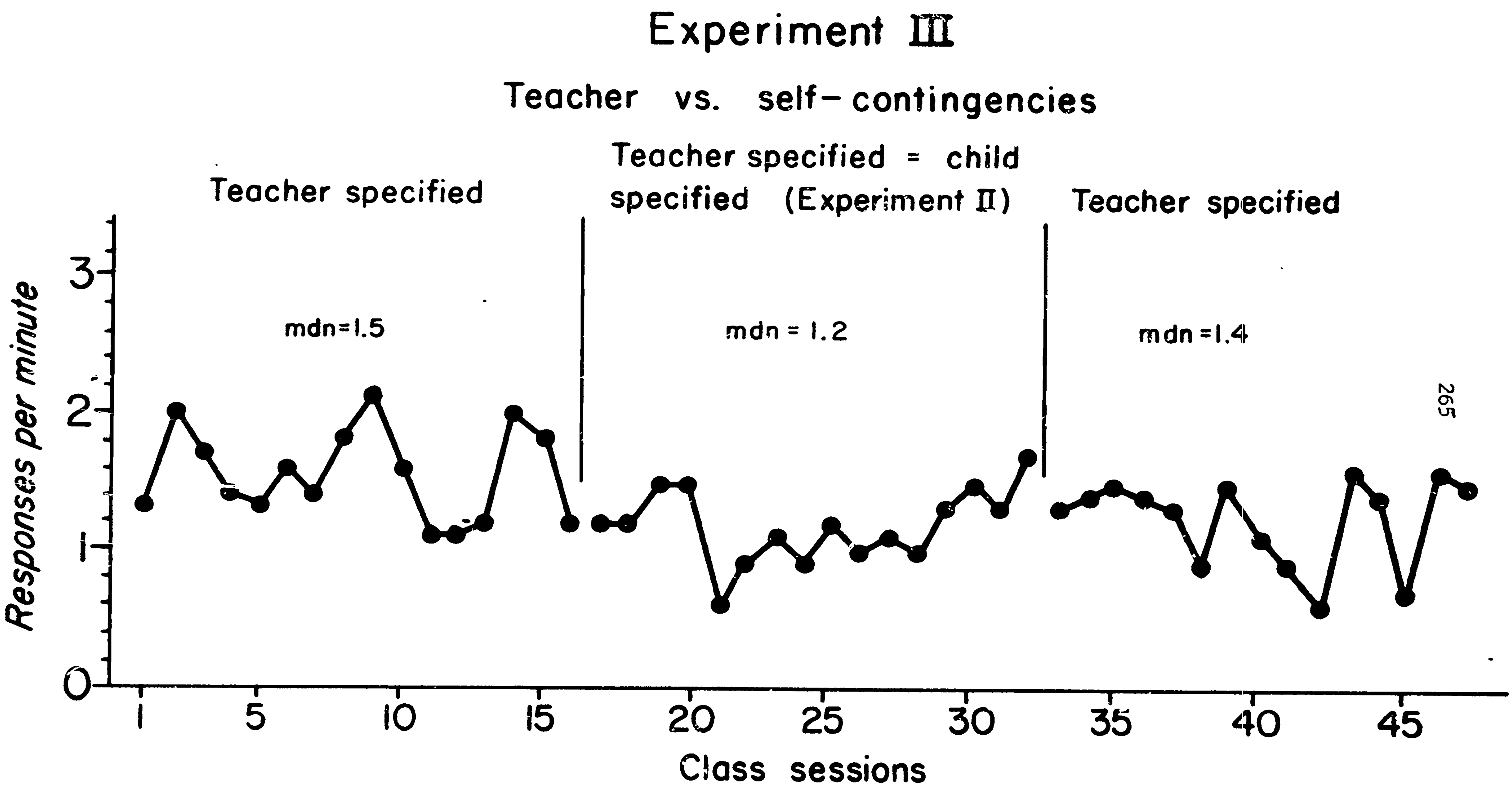
Teacher vs. self - contingencies



Experiment II

Teacher vs. self-contingencies




Experiment III
Teacher vs. self-contingencies
Teacher specified
Teacher specified = child specified (Experiment II)
Teacher specified
mdn=1.5
mdn = 1.2
mdn =1.4
Responses per minute
0
1
2
3
1
5
10
15
20
25
30
35
40
45
Class sessions
265

subject's median response rate was 1.5 for the 16 days of Stage 1. During Stage 2, which also lasted for 16 days, his median response rate was 1.2; and during the final stage, consisting of 15 sessions, it was 1.4. The range during Stage 1 was from 1.1 to 2.1; from .6 to 1.7 during Stage 2; and from .6 to 1.6 during Stage 3.

Conclusion

The data from these experiments indicated that, for this subject, self-imposed contingencies were associated with an increased academic response rate. This was evidenced in Experiments I and II, for during Stage 2 of each experiment, the period of self-contingencies, the subject's median performance was higher than during Stages 1 and 3, the periods of teacher-imposed contingencies.

That this response rate increase was attributed to the manipulation of the contingency manager and not to the contingency system being explained or unexplained, was demonstrated in Experiment I. Although the latter manipulation might alter the response rate of some subjects, the data relevant to this variable indicated that, for this subject, explaining or not explaining the contingencies produced nearly equal effects (see Figure 1).

In addition, the data from Experiment III revealed that the response rate increase was due to manipulation of the contingency manager, and not to reinforcement magnitude. As mentioned earlier, although two experiments had demonstrated that during periods of self-specified contingencies the subject responded at rates higher than during periods of teacher-specified contingencies, this increased rate may have been due to mere magnitude of reinforcement. Indeed the contrary seemed to be the case, for when the reinforcement ratios were increased by the teacher during Stage 2 of Experiment III, the subject's response rate decreased, while it increased during Stage 3 when the original payment was in effect.

Throughout this study response units from the various areas of reading, math, and spelling were grouped under a single category, "academic response." It may be, therefore, that the effects of the subject's specifying his own contingencies were more pronounced in certain academic areas than in others. It is also possible that when a subject is allowed to specify his contingencies in one area but not in another, he would, in fact, respond at an accelerated rate in both areas. Research is presently being conducted, therefore, in which the effects of self-imposed contingencies are being analyzed within separate areas of academic responding in order to specify more precisely the effects of self-contingency management.

The evidence from this current experiment, that self-scheduling of events is associated with accelerated performance, is supported, in part, by a recent experiment (Lovitt & Curtiss, 1968). In order to investigate the effects of choice as an independent variable, subjects were scheduled a number of daily sessions, each comprising three phases. One phase consisted of assigned mathematics, one of assigned reading, and one of a choice period in which the subjects could select either reading or mathematics. The results revealed that the subjects' rates of responding were greater during choice periods than during no-choice periods. It appeared that, for the subjects in this study, to choose (even between two academic tasks) was a reinforcing event.

If continued explorations with self-contingency management reveal similar findings, the educational implications appear rather obvious; for not only does the individual begin to develop self-managing skills by arranging certain aspects of his own environment, but in so doing his academic performance increases.

Research is now in progress to specify other components of self-management. Once these elements have been detailed, not only can they be programed sequentially, but each can be independently manipulated to evaluate its effect on academic responding. One study currently being conducted compares differences in student performance when a child has a graph available showing his daily performance rates and when he does not have such a graph. A second investigation is designed to evaluate the effects on academic performance when a student plots his own performance data vs. the effects when the teacher does the plotting. Another study is concerned with the function of specifying academic requirements. During certain phases of this study, the teacher sets the program limits while during other portions of the study the student specifies the limits for daily performance.

Self-management no doubt involves behaviors other than those specified in this report. The fact remains, however, that if in education we are committed to educate students so that they can not only discriminate a number of teacher-arranged stimuli, but can also arrange their own environment--hence control their own behavior--we must conceive of these self-controlling behaviors as behaviors that can be taught and learned. The behaviors leading toward self-management, therefore, must be independently investigated and sequentially arranged to formulate a self-management curriculum.

Study II: Rate-Contingent Reinforcement: A Tactic to Increase Academic Response Rate

That various contingencies of reinforcement affect an organism's behavior in a predictable manner has been widely demonstrated (Ferster and Skinner, 1957). Morse (1966), in fact, has stated that these contingencies of reinforcement are as influential in generating and maintaining behavioral patterns as the reinforcers themselves. He further commented that "powerful control of behavior by discriminative stimuli and by reinforcers such as food and water actually develops because they are favorably scheduled events."

Field situations, however, have been more concerned with attempts to locate reinforcers than with a parametric analysis of the contingencies of reinforcement. While it is essential to ascertain the events in a person's environment that may be used to maintain his rate of appropriate responding in academic settings, the manipulation of contingencies or the temporal arrangement of events must also be considered by the therapist in his efforts to assist the pupils to acquire and maintain behaviors.

The investigation reported here, which comprises a series of experiments, was prompted by a boy who responded academically at a very low rate. A previous attempt to accelerate this subject's rate of response involved the manipulation of reinforcers by changing the consequence of his academic behavior from contingent time with games and crafts to social time with an adult male. This manipulation did not significantly affect the dependent variable, academic rate.

The current experiments sought to analyze the function of a contingency manipulation on response rate. The purpose was to compare a subject's response rate when a single fixed ratio contingency was in effect (a specified number of points for a specified number of responses) with his response rate when the contingencies (points granted in direct proportion to his rate of responding) were in effect.

Experiment I

Method

The subject was a 12-year-old boy enrolled in a class for children with behavioral disorders at the Experimental Education Unit of the University of Washington. The material used during this experiment was the subject's regular mathematics material, Sets and Numbers, by Suppes (Singer Mathematics Program, 1965).

During the 15-day baseline phase of this study, data were obtained one hour each day on the subject's response rate to the math material. At this time the subject was on a 20:1 schedule--one minute of free time in the "high-interest" room (Haring & Lovitt, 1967) contingent on 20 correct mathematics responses.

During the second phase of this study (33 trials), differential rate contingencies were initiated. These new specifications were derived on the basis of the subject's performance during the first (20:1) phase. In order to receive any pay-off during Phase II (differential rate condition), the subject had to respond beyond his Phase I rate of about I per minute. Response rates higher than the initial performance were reinforced at correspondingly richer ratios. The four contingency bands were as follows: (1) no points if less than 60 responses were emitted; (2) three points for 60-89 responses; (3) nine points for 90-119 responses; and (4) 15 points for 120 or more responses. The 60th, 90th, and 120th responses were marked on the subject's math sheet. These markings and the new differential contingency system were explained to the subject each day. Table I describes the 4 differential contingencies in terms of response requirements, rate equivalent, points earned, and the point-per-response or ratio equivalent.

In phase 3, extending over 7 trials, the initial contract was reinstated--one point or minute of time per 20 mathematics responses. Throughout the experiment the teacher calculated the subject's rate immediately following each session. Then, dependent on his correct response rate, the teacher would give him the corresponding number of points, which could then be redeemed for minutes of free time.

Results

During the beginning phase of this study the subject's median response rate was .8, while his response rates ranged from 0 to 2.9. His median mathematics response rate during the differential reinforcement stage was 1.7. His range of responding throughout this period was 3.8, extending from

Table I

Differential Rate Contingencies

Experiment I

Responses	Rate/Minute	Points	Ratio
<60	<1	0	0
60-89	>1	3	20:1
90-119	>1.5	9	10:1
>120	>2	15	8:1

.2 to 4.0 responses per minute. A median response rate of .6 was calculated for the seven day period where the initial, premanipulation conditions prevailed. The subject's range during this period was 1.3, extending from .2 to 1.5 problems per minute. Figure 1 presents the daily response rates throughout the experimental sessions.

Although performance appeared to be sensitive to the experimental manipulation of the variable, differential-rate contingencies, it was decided to run a second, more carefully controlled experiment. The basis for this decision was that although the subject's overall rate of responding did increase when differential contingencies were in effect and subsequently decreased whe the variable was removed, the median difference from condition to condition was not great and the subject's responding during all experimental phases fluctuated widely.

Experiment II

Method

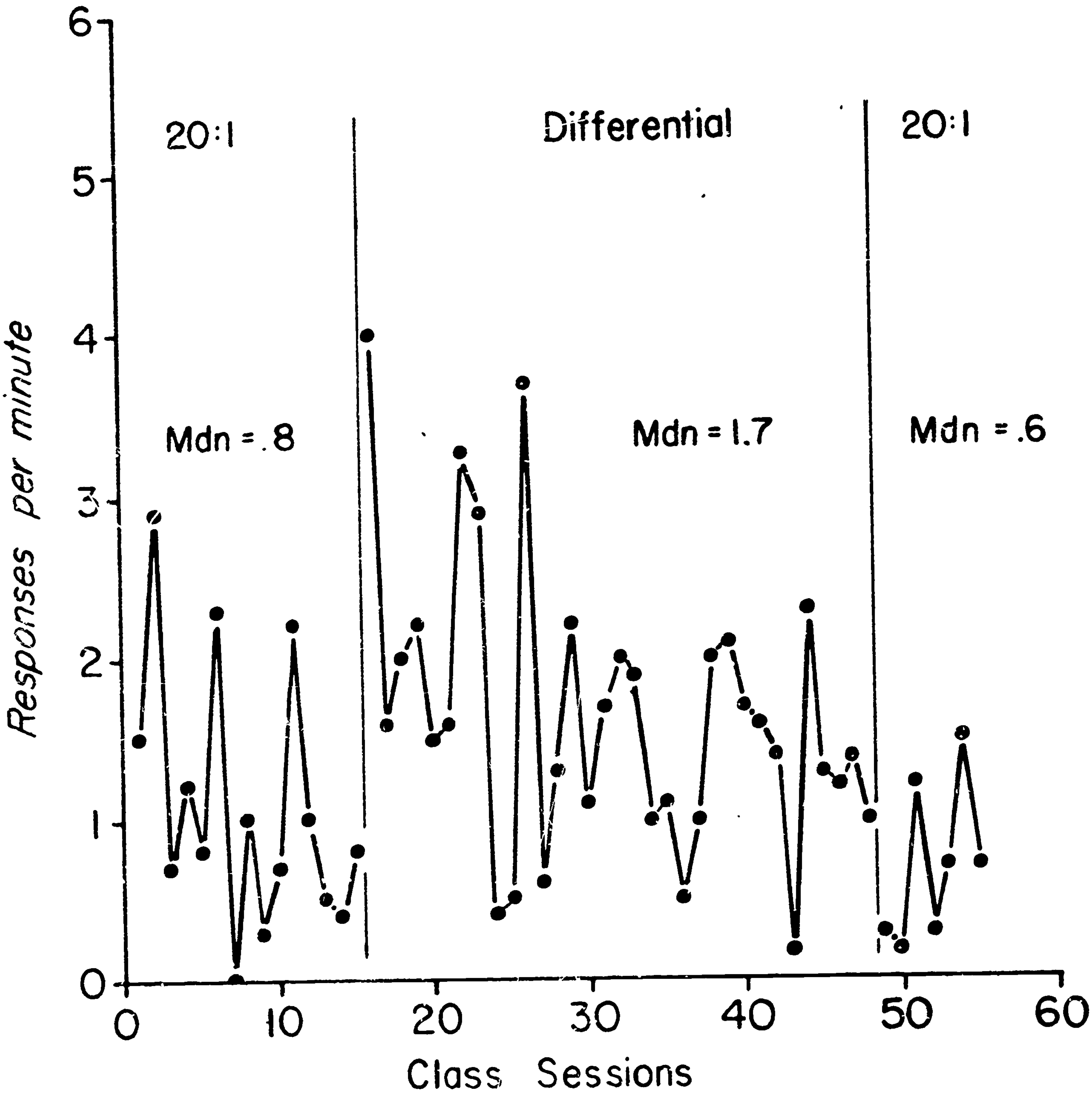
Since the major concern of this investigation was to assess the variable of differential rate requirements and the subject's acquisition of mathematics responses was only of secondary concern, the academic material was altered in this experiment. Rather than require the subject to respond to mathematics material from the Suppes program, as during Experiment I, the subject was now given mathematics problems of the class 49 + 23=_______, where the sum was $\underline{\leq}$ 198. Mathematics problems of this class were in the subject's repertoire (Easy Math Program).

During phase 1 of this study (12 trials), the subject's point-per response requirement was 20:1 (the same ratio that was in effect during phases 1 and 3 of Experiment I). In phase 2, consisting of 12 trials, differential rate contingencies were again imposed. As during Experiment I, the 4 contingency bands were derived on the basis of the subject's median performance in the first phase of the previous experiment. Since the subject's median rate was about 3 responses per minute during the initial phase of Experiment II, he had to respond beyond that rate to receive points during the differential pay-off phase. Response rates below 3 per minute were not reinforced though rates beyond 3 were reinforced with successively richer ratios. The 4 contingency bands specified the following rate requirements: (1) no points for 0-44 responses; (2) 3 points for 45-59 responses; (3) 6 points for 60-74

EXPERIMENT I

Differential Rate Contingencies

(Suppes Math Program)



responses; and (4) 15 points for 75 or more responses. Table 2 details the differential contingencies as to required responses, rate equivalent, points earned, and ratio equivalent.

Should the subject respond at a rate beyond 4 problems per minute, his pay-off would be at the rate of 10:1. And should he emit 75 or more responses during the 15-minute session (a rate beyond 5 per minute) his pay-off would be calculated on a 5:1 basis. The 45th, 60th, and 75th mathematics problems were marked on the subject's work sheet, and, as in Experiment I, the differential aspects of the contract were explained daily to the subject.

During the third phase, which was composed of 8 trials, the original 20:1 contingencies were in effect. In the fourth phase of this study, comprising 8 trials, the differential rate contingencies were once again in effect.

Results

During the first phase of Experiment II the subject's median rate of response was 3.1, ranging from 2 to 4.5. A median rate of 5.35 responses per minute was obtained in the second phase, varying from 4.0 to 7.3 responses per minute. A median response rate of 3.9 was calculated for the third phase, from 1.7 to 4.4. A median rate of 6.4 was obtained for the final, rate-contingent phase. The subject's response throughout this phase varied from 5.5 to 8.4. The data from 4 phases of this second experiment are presented in Figure 2. The ranges for the 4 experimental conditions were 2.5, 3.3, 2.7, and 2.9. This variability was virtually the same as that reported during Experiment I.

Experiment III

Method

Although the variable, differential rate reinforcement, apparently was effective in altering the subject's rate of mathematics responding, the possibility existed that sheer magnitude of reinforcement had been at least partially responsible for the performance increase. As can be noted from Figure 3, the subject received much more reinforcement in terms of the number of points during the rate-contingent phases of Experiment II than during other phases of the study. In fact, during the experiment when a comparison is made between the subject's response rate and the rate at which points were dispensed, the patterns were very similar, i.e., when responses per minute were high, the number of points received was also high (See Figures 2 and 3).

Table 2

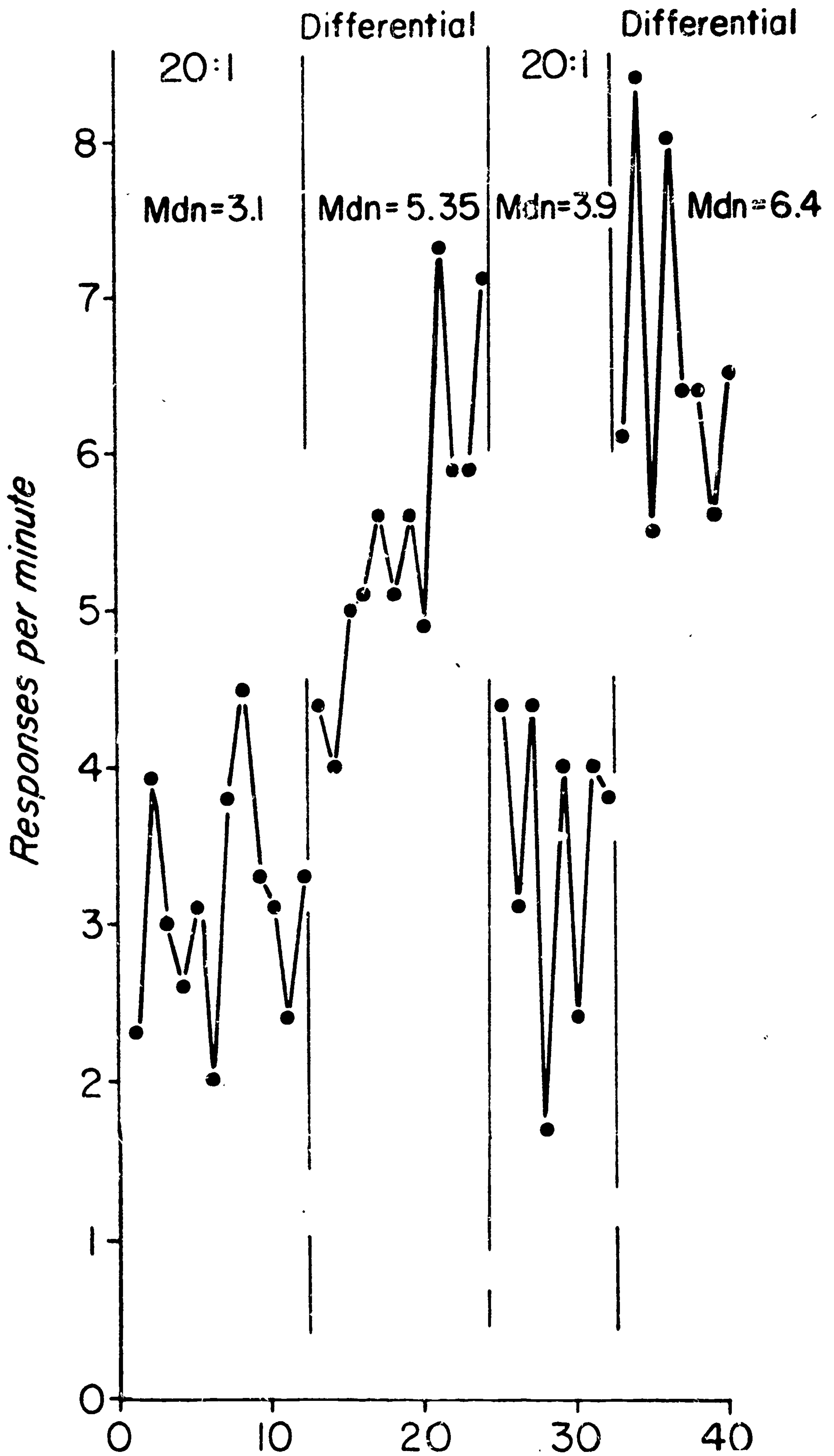
Differential Rate Contingencies

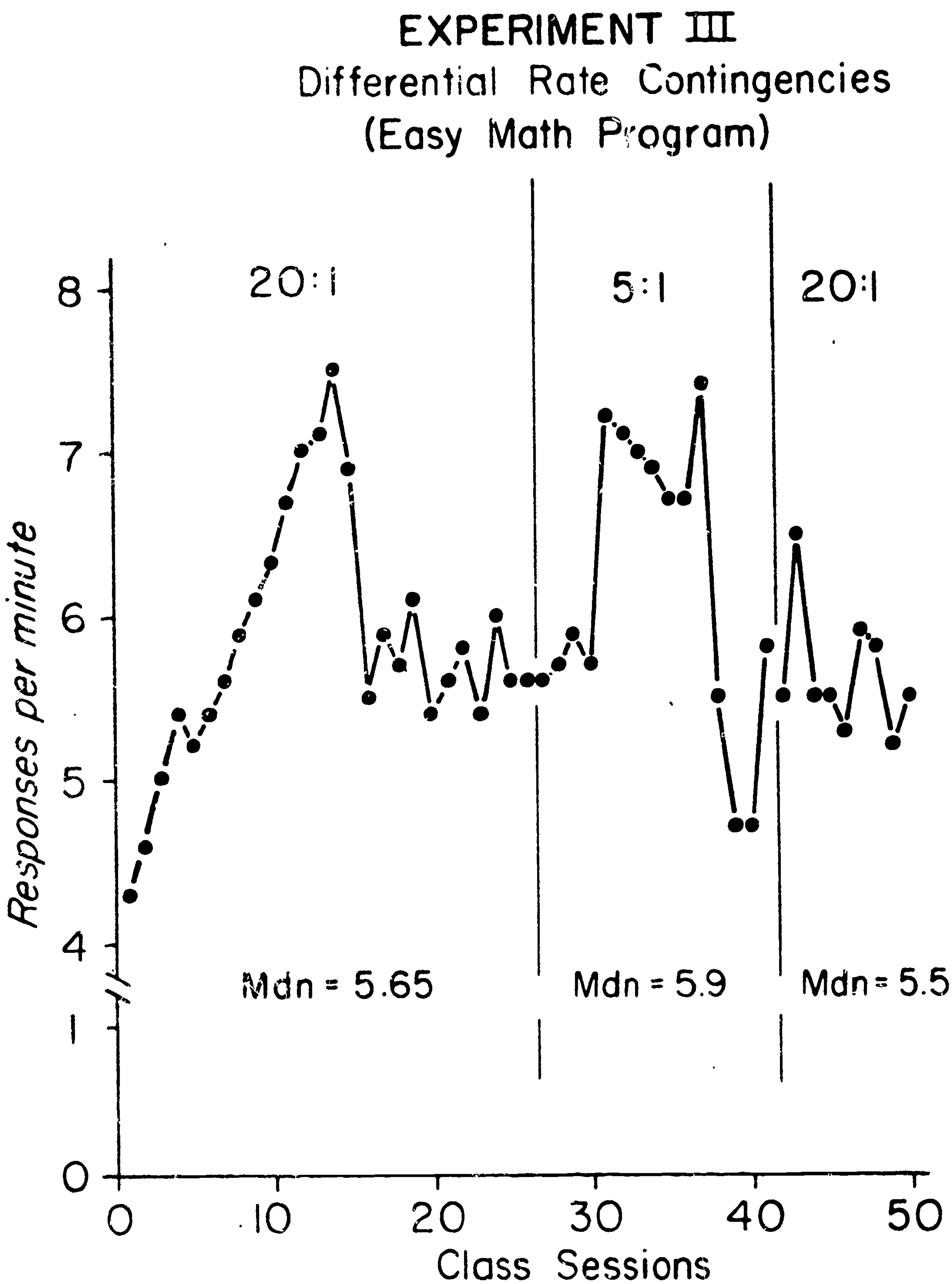
Experiment II

Responses	Rate/Minute	Points	Ratio
< 44	< 3	0	0
45-59	> 3	3	15:1
60-74	> 5	6	10:1
> 75	> 5	15	5:1

EXPERIMENT II

Differential Rate Contingencies
(Easy Math Program)




EXPERIMENT III
Differential Rate Contingencies
(Easy Math Program)
20:1
5:1
20:1
Responses per minute
8
7
6
5
4
1
0
Mdn = 5.65
Mdn = 5.9
Mdn = 5.5
0
10
20
30
40
50
Class Sessions

The procedures of this experiment were the same as those during Experiment II. The subject was reinforced on 20:1 ratio during phases 1 and 3, while differential rates were in effect in the second and fourth phases. The type of mathematics problem and the length of each experimental session were also the same as before - "easy" materials and 15-minute sessions.

The differential ratios employed in phases 2 and 4 were the same as during Experiments I and II. These rates were based on the subject's average response rate during the initial 20:1 phase, which during Experiment IV was 6 responses per minute. In a 15-minute session, the subject would be expected to emit 90 responses (15x6). Therefore, the four differential rate ratios were: (1) no points for 0-90 responses; (2) 6 points for 90-99 responses; (3) 8 points for 100-119 responses; and (4) 12 points for 120 or more responses. Table 3 presents the differential contingencies as to required responses, rate equivalent, points earned and ratio equivalent.

Results

Throughout Experiment IV it was apparent that the cue marks used during the differential phases of the study were not exclusively responsible for the accelerated performance during differential rate phases. This variable, cue marks, was not singularly manipulated but was incorporated throughout all phases of Experiment IV. Therefore, the generally increased rates from Experiment II to Experiment IV could have been due in part to the marks. However, the differential rate contingencies continued to be an extremely effective variable, as can be noted by comparing the differential phases with 20:1 phases in Figure 5.

During the first phase of Experiment IV, the uniform contingency phase, the subject's responding ranged from 3.9 to 7.4, a median response rate of 6.1. Throughout the 15 sessions of phase 2 his median rate of responding was 8.2, ranging from 7.5 to 9.6.

When a reversal of contingencies was initiated, the subject's responding ranged from 5.5 to 7.8. His median rate throughout this phase was 6.8. The fourth phase of Experiment IV, the return to differential rate contingencies, was characterized by a response range of 1.9 (7.6 - 9.5) and a median rate of responding of 8.4.

Table 3

Differential Rate Contingencies

Experiment IV

Responses	Rate/Minute	Points	Ratio
< 90	< 6	0	0
90-99	> 6	6	15:1
100-119	> 6.6	8	12.5:1
> 120	> 8	12	10:1

To determine, then, whether the rate-contingent variable or magnitude of reinforcement had been the crucial variable in affecting response rate differences, magnitude of reinforcement was the only variable manipulated during Experiment III. The first and third phases of the experiment, extending over a period of 50 sessions, employed the same base ratio, 20 responses received I minute of free time. A 5:I ratio, the highest possible during the previous rate-contingency phases, was used during the second phase of the study. The arithmetic materials were the same as those used during the second experiment (problems in the subject's repertory).

Results

The results of Experiment III, illustrated in Figure 4, revealed that apparently magnitude of reinforcement was only a minimally affecting variable. The subject's response rate ranges were 3.2 (4.3-7.5), 2.7 (4.7-7.4), and 1.3 (5.2-6.5) during the first, second, and third phases respectively. His median rates of response were 5.65 during phase I, 5.9 during phase 2, and 5.5 during phase 3.

Experiment IV

Method

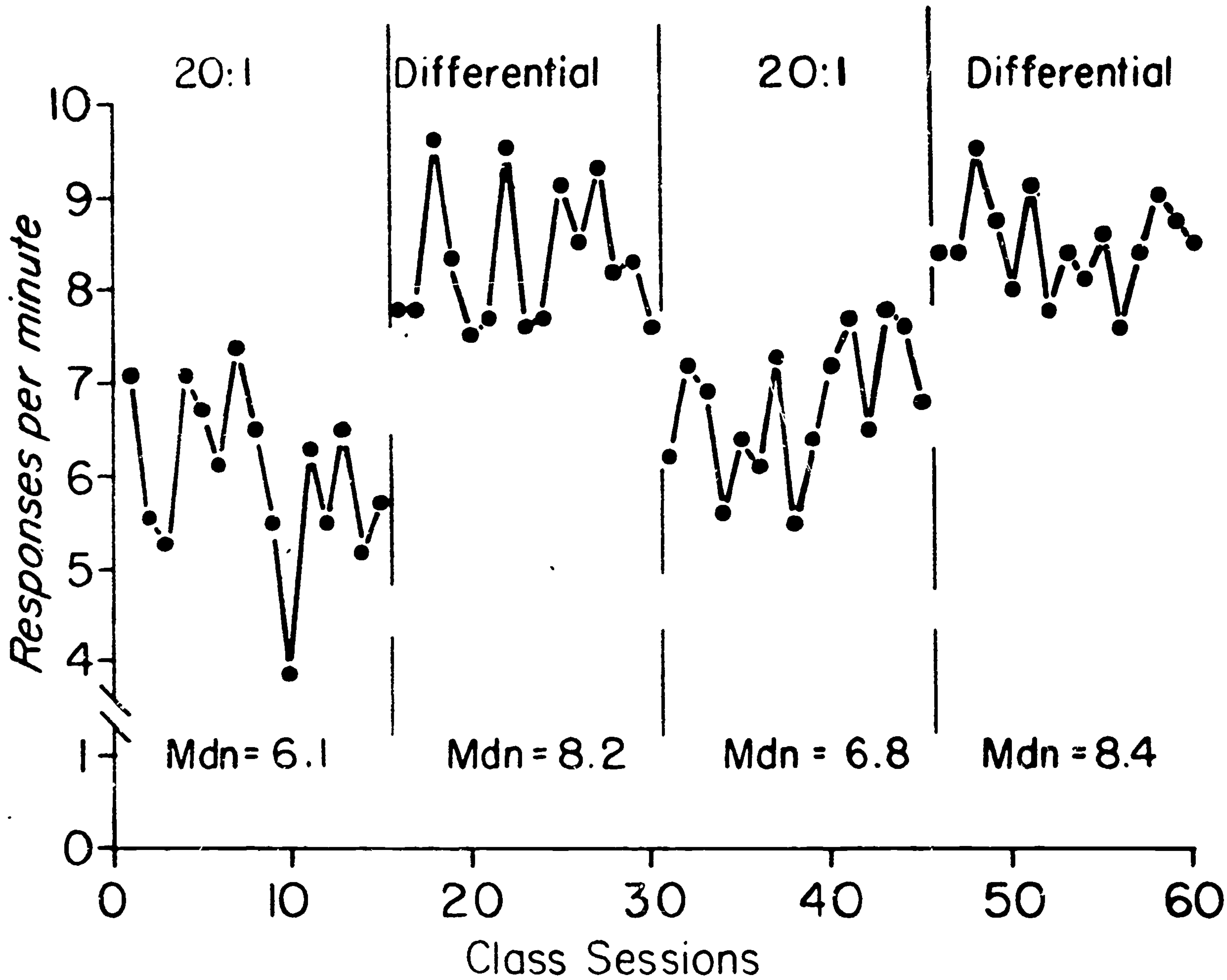
Experiment IV was conducted for two reasons. First, as noted from the data in Figure 4, the subject's rate of responding was to some extent affected by magnitude of reinforcement. The effects of this variable were, however, rather insignificant, particularly when compared to the effects of differential rate as imposed during Experiment II. Nevertheless, since some change resulted when magnitude of reinforcement was altered, it was decided to once again manipulate differential rate contingencies. By replicating Experiment II, the investigators would be able to evaluate more definitively the relative effects of the two variables, since experiments manipulating differential rates would precede and follow Experiment III, which dealt with reinforcement magnitude.

A second reason for conducting the fourth experiment was to determine whether the marking of certain problems to indicate performance rate was of itself accountable for the subject's rate increase. During the differential reinforcement phases of Experiment II, the 45th, 60th and 75th responses were marked on the subject's math sheets to indicate which differential reinforcement would prevail. These marks were therefore included throughout all phases of Experiment IV.

EXPERIMENT IV

Differential Rate Contingencies

(Easy Math Program)



Discussion

Several tentative observations may be made from the studies contained in this investigation. First, differential contingencies of reinforcement were more effective in accelerating academic performance than sheer magnitude of reinforcement. This observation was illustrated in a comparison of experiments III and IV. For although the same materials were used during those two experiments, the dependent variable was affected much more in Experiment IV where rates were manipulated than in Experiment III where reinforcement magnitude was altered. This observation of the influential aspects of reinforcement contingencies supports the earlier cited claims of Morse.

Another observation, noted when the data from the first differential rate phase of Experiments I, II, and IV were analyzed, was that the higher rate requirements generated more pronounced behavioral effects than the lower rates. For when subject responding was analyzed during the 4 ratios of the differential rate segments, it was discovered that the subject was paid off more often from the highest rate contingency than during any of the others. During phase 2 of Experiment I the subject's responding fell 5 times within the no pay-off band, 10 times in the next to lowest, seven in the next higher, and 11 in the highest. During phase 2 of Experiment II, his rates fell 3 times in the next to highest and 9 times in the highest bands, while never falling in the no pay-off or next to lowest rate bands. In Experiment IV the subject was also paid off from only the top two bands--7 times from the next to highest and 8 from the highest.

On the basis of this series of studies, it appears that differential rate contingencies should be considered as an affecting variable when the objective is to accelerate academic response rate. Furthermore, consideration should be given to the manipulation of ratio or interval contingencies rather than only to finding more powerful types of consequences.

This series of experiments also showed that when the curriculum was stabilized by using math materials which were all of the same class and within the response capabilities of the subject, the effects of the independent variable, differential rate contingencies, were more readily discernible than when math problems of a variety of forms and levels of difficulty were used. The use of uniform materials could prove a useful tactic when investigators are more concenred with the effects of an imposed independent variable on the dependent variable, academic response rate, than with the acquisition of academic behaviors.

Study III: A Functional Assessment of Teaching and Teacher Training

Much debate has arisen regarding the preparation of teachers. Many discussions have been concerned with the practicum experience--how much should be provided, and at which part of the training sequence it should appear. Other discussion has focused on the course content of the training sequence--whether more time should be allotted to courses of subject matter or to teaching methodology. Recently, too, some critics of teacher preparation programs have voiced concern over the lack of liberal arts exposure in the training of some prospective teachers.

However, few, if any, of the supporters of certain training styles can offer empirical evidence to support their preferences. The obvious reason for this lack of data relevant to preparation sequences is that no dependent variable exists. In other words, the "effective" teacher has not been defined. Until this is accomplished, efforts to produce such a person are apt to be fruitless.

Although various attempts have been made to define teachers as those who participate in community affairs, are accepted by their peers, have taken certain requisite courses, hold recommendations from notable citizens, or have scored high on standardized tests, the function of such information is dubious. It is quite possible that some individuals who rank high in regard to these items may prove to be ineffective teachers. Conversely, certain individuals who rank very low on the same items are, in fact, excellent instructors.

Since the goal of the teacher is to modify behavior--to assist children to acquire, to arrange, and to use information--teachers must be evaluated on the basis of pupil change, on the number of mathematics, social studies, reading, writing, and science skills they can teach pupils. Furthermore, they must be evaluated as to the rate at which they teach certain skills, whether in one day or one year.

Teacher training institutions must be similarly evaluated on the number and rate of skills that their students are able to teach children. A teacher training institution preparing students to teach 200 skills in one year would be inferior to the institution whose students could teach 2000 skills in the same time.

In the current study the paradigm of teacher trainer, student instructor, and pupil was simulated by a classroom teacher and three of her students. Of the three boys selected from a class for children with behavioral disorders at the Experimental Education Unit (EEU), University of Washington, one student acted as the student instructor, while the others served as pupils. The three boys, ages 10-12, had been members of the EEU classroom for 2-3 years. As members of this class, the pupils individually are granted points redeemable for free time activities contingent or correctly answering a certain number of academic problems. These point-per-answer contingencies are individually arranged, varying from one pupil to another and from one academic area to another.

The design of this series of studies was to specify certain teaching procedures, to teach them to an individual, and to arrange a setting whereby he might use the procedures. Furthermore, a cooperative contingency system was established to reinforce both learning and teaching. Contingencies were arranged so that both the student instructor and pupil were provided reinforcement dependent on the performance of the pupil. The objectives of the studies were to determine student instructor efficiency (number of skills taught the pupil as a function of the instructor's time) and teacher training efficiency (number of skills taught the pupil as a function of the trainer's time).

Experiment I

Procedures

During the teaching process two boys were seated at a table on one side of the room. The student instructor was given a set of 20 words and asked to teach these to a fellow pupil. He was also told to use any teaching method he preferred for a drill period lasting no longer than 10 minutes, though he could, if he wished, terminate the session earlier. The purpose at this time was to determine how successfully the student instructor could teach without any procedural assistance.

The 20 words, selected from Book 13 of the Sullivan reading series, were words the pupil was unable to pronounce. An attempt was made to select words of varied lengths and letter combinations to avoid, as much as possible, generalization from one word to the next. Words such as "dart," "climb," "branch," and "thirty" were written on 3x5 cards and provided the student teacher.

After the drill session, the classroom teacher (teacher trainer) tested the pupil, asking him to name each of the 20 cards that had been used in the drill sequence. No confirmation was given the pupil for either correct or incorrect responding. The teacher trainer timed the drill session of the student instructor and the pupil, recorded each correct and error response during the posttest, and later translated these figures into correct and error rates (number of correct and incorrect responses on the posttest, each divided by the time of the drill session).

At the beginning of this study the teaching pair was told that payment would be made on the basis of the pupil's performance on the posttest. A 5-to-1 contingency ratio; both the pupil and the student instructor received one point for every 5 correct pupil responses. This study was run until the pupil's error rate was zero for two consecutive days during the posttest.

Results

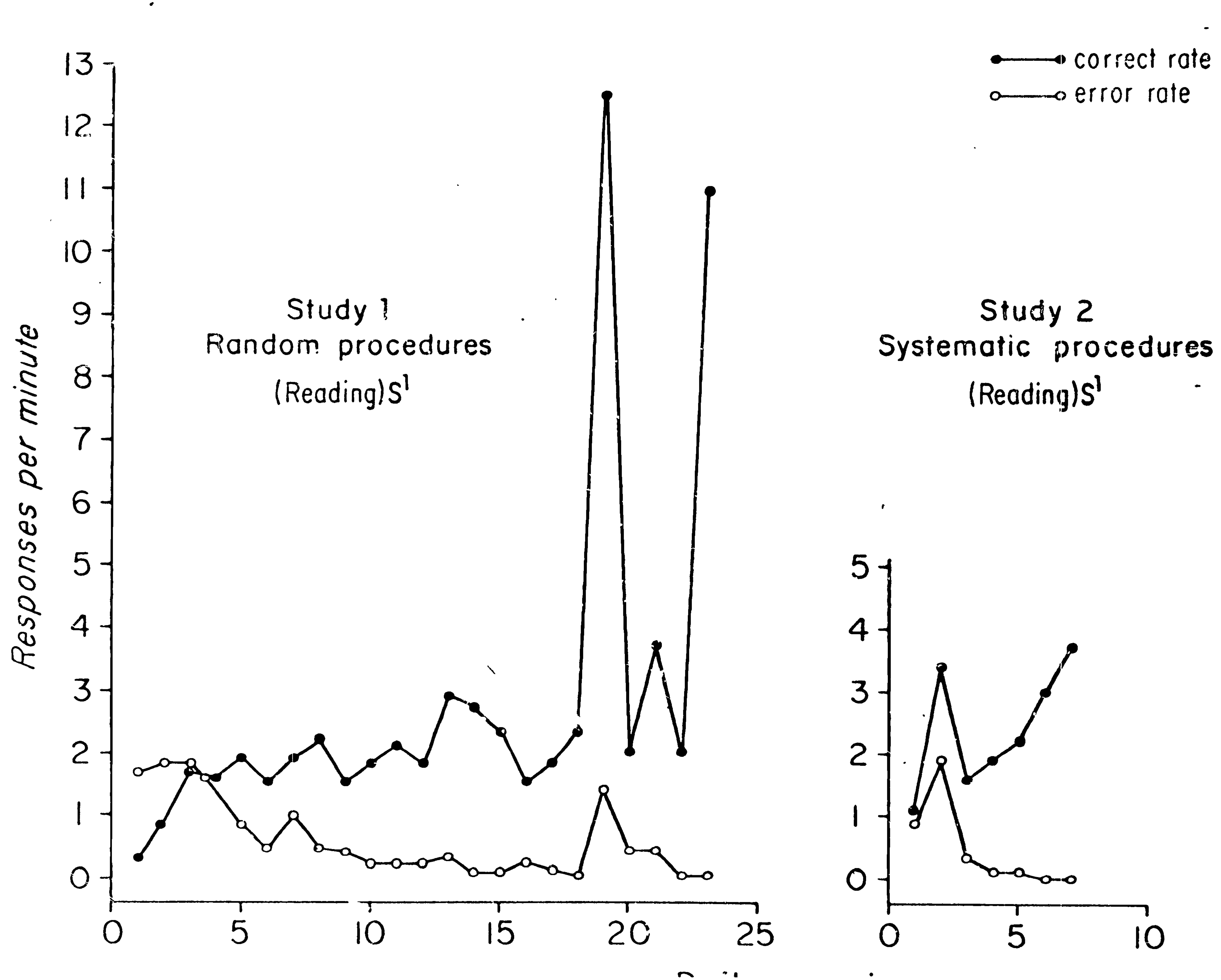
Following the first day's teaching session the pupil made 17 errors, answering correctly only "shovel," "brother," and "body." The next day he answered 7 correctly, the third day, 10, and the fourth, 14. By the tenth day he answered 18 correctly; and finally, on the 22nd and 23rd days this performance reached the criterion set for him: no errors. From the 10th to the final session, two words, "highness" and "another" remained troublesome for the pupil.

As indicated by Figure 1, the pupil's correct rate gradually progressed while his error rate gradually fell. Figure 1 also shows that on two days the pupil obtained extremely high correct rates: 12.5 and 11.

Experiment II

Procedures

The purpose of the second experiment was to determine the effects on pupil performance when certain instructional procedures were used by the student instructor. An additional twenty words not familiar to the pupil, such as "excitement," "robbers," and "elevator" were selected by the teacher trainer from Book 14 of the Sullivan readers. Again, the pupil and student instructor were allowed up to 10 minutes for the drill period and were pro-



correct rate
error rate
Study 1
Random procedures
(Reading)S1
Study 2
Systematic procedures
(Reading)S1
Responses per minute
13
12
11
10
9
8
7
6
5
4
3
2
1
0
5
10
15
20
25

vided space at a table in the classroom. Now, however, unlike Experiment I, the student instructor was furnished with some specific teaching instructions. He was told, "Show the pupil one card at a time and ask him to name the word. If he answers correctly, say 'good' and go on to the next card. If he answers incorrectly, say the word and then have him repeat it.' Following this training session, the teacher trainer tested the pupil and calculated his correct and error rates. As in Experiment I, the rate was derived by dividing the number of correct answers or errors by the time of the drill session.

Results

When the teacher trainer administered the test after the first training session, the pupil correctly answered 11 of the 20 words. All items were correctly answered following the sixth and seventh training sessions. The pupil's rate of correct answers, indicated by Figure I, gradually increased from the first to the seventh session, while his error rates correspondingly decreased. Further, the figure shows that only 7 training sessions were required to attain criterion performance when teaching procedures were given the instructor, whereas 23 were needed when no teaching procedures were used.

Experiment III

Procedures

In Experiment II the teacher trainer computed the rates, graphed the data, and administered the posttest. The purpose of Experiment III was to teach the student instructor to perform these additional teaching procedures.

Throughout this study, demonstration, imitation, and confirmation were employed by the teacher trainer. First, the teacher trainer demonstrated each successive teaching skill to the student instructor. The roles were then reversed, the student instructor imitating what had been demonstrated and the teacher trainer acting as the pupil. The teacher trainer provided verbal confirmation to the student instructor immediately following his efforts to imitate the teaching procedure. The criterion for successful performance throughout the teaching of these procedures was two consecutive correct performances. If the student instructor failed in any one step, the teacher trainer repeated that phase of the teaching; the student instructor was advanced to the next step only when his competence on the previous step was evident. Throughout the teaching of these steps, the teacher trainer recorded both the number of teaching trials and the amount of time required to teach them.

The student instructor was first "taught all the behaviors necessary for the drill segment of the session--to indicate the session number on the data record sheet, and to start and stop the stop watch at the beginning and end of each drill session. The student instructor was then retested on the teaching procedures he had used during Experiment II--the presentation of the cards and the confirmation of pupil responses. Next he was taught to calculate the total time of the drill session, to convert this figure to seconds, and to record the time on the data sheet.

When the student had acquired the skills necessary for drilling, he was taught the behaviors needed to test the effectiveness of the drill, which originally had been handled by the classroom teacher. The first skill required scrambling the 20 sight words in order to eliminate the possibility of the pupil's answering on a chaining or associative basis. Second, he was shown how to tally the pupil's errors and correct answers. On completion of the test session, the student was taught to total both the number of errors and correct answers, to record these figures on the data sheet, and to compute the points earned by his pupil (1 point per 5 correct answers on the test). He was then instructed on procedures for dismissinh his pupil when the training and test procedures were completed.

The student instructor was also shown how to calculate the rate by dividing total error and total correct responses on the posttest by the time in seconds that had elapsed during the drill, then told to multiply these rates by 60 to obtain a rate-per-minute statement. Finally, the student instructor was shown how to record these data on a record sheet and to graph daily error and correct response rates. The teacher trainer spent 289 minutes in teaching the student instructor these procedures.

Experiment IV

Procedures

During this experiment the original dyad, student instructor and pupil, again participated. The purpose was to determine which procedures the student instructor would use and what effect, if any, would be evidenced by his pupil since now the pupil was involved with only the student instructor, not with the student instructor and the teacher trainer. Twenty unrelated words were selected from the Sullivan Book 15, such as "repeat," "crazy," "subject," and "director." As in the first two studies, the classroom teacher pretested the pupil and verified that these words were unknown.

The basic procedures throughout this study were similar to those used in the first and second studies: a drill session of no longer than 10 minutes followed by a testing session. The student instructor was given sole responsibility for the project, including training, testing, and recording, using the procedures he was taught during Experiments II and III.

Results

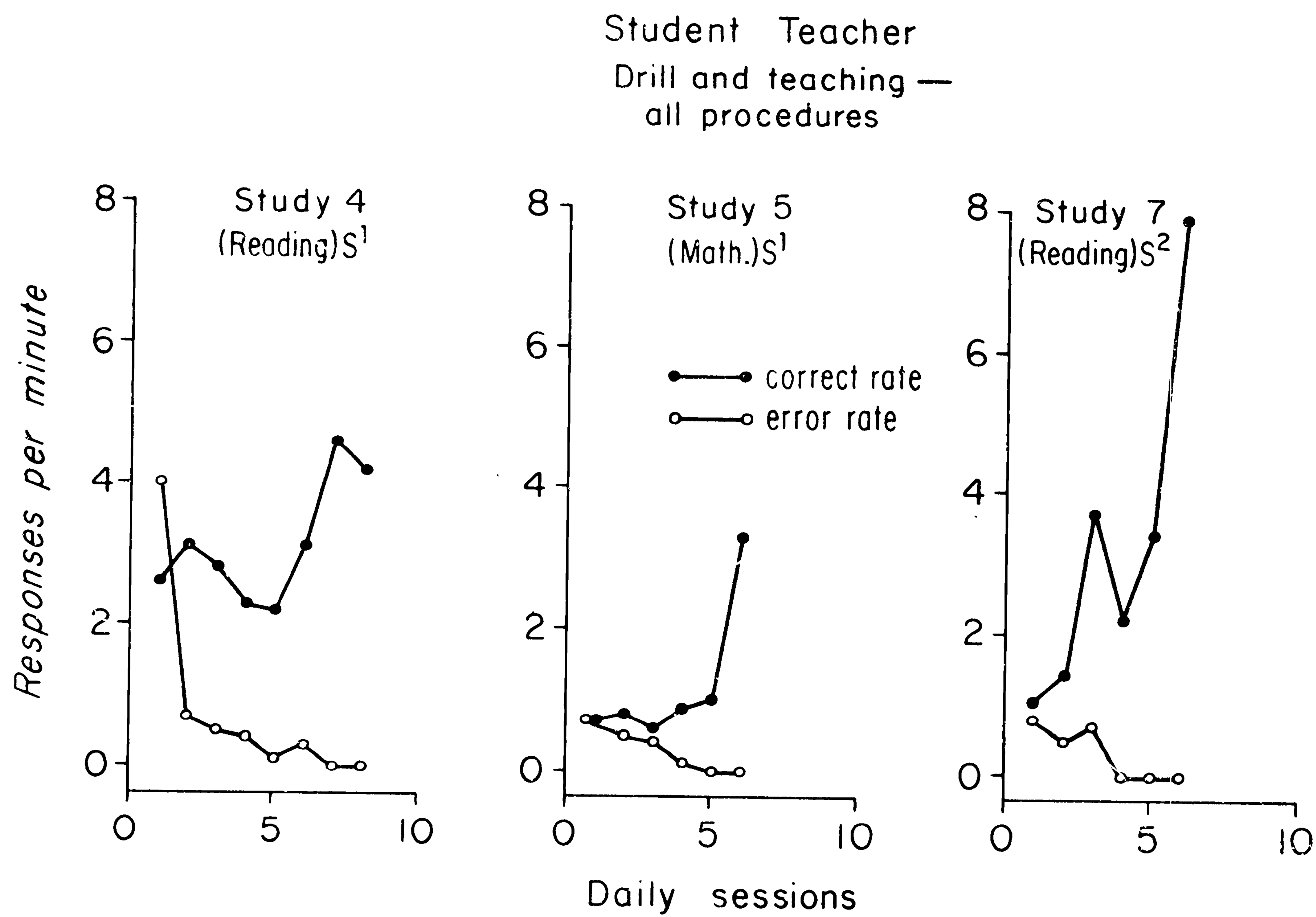
The data relevant to the pupil's correct and error rate performance are shown in Figure 2. As indicated, 8 sessions were required for the pupil performance to reach the criterion of 2 consecutive no-error sessions. While the student instructor conducted all the procedures of this study, the classroom teacher observed, noting any errors or omissions in the teaching procedures. Only two such procedural errors were noted: one, when the student instructor did not start the stop watch at the beginning of a testing situation, the other, when he did not confirm the pupil's correct response. In both instances the student instructor was reminded to perform these acts.

Experiment V

Procedures

The same two boys were used during Experiment V. However, now the student instructor's task was to teach his pupil a set of 10 multiplication problems instead of 20 words. These multiplication problems were of the class 7x3=___and 3x8=____. Throughout these sessions the student instructor was told to use the teaching procedures of his choice.

The purposes of this study were to determine how many of the teaching procedures programed during the preceding studies would generalize to another subject matter area, and how effective these procedures would be in another area. In order to evaluate the student instructor's competence, two dependent variables were obtained daily. First, the performance rate of the pupil was obtained in terms of error and correct response rates. The second dependent variable was the number of appropriate teaching techniques used by the student instructor. To obtain this latter data, the


Student Teacher
Drill and teaching — all procedures
Study 4
(Reading)S1
Study 5
(Math.)S1
Study 7
(Reading)S2
correct rate
error rate
Responses per minute
Daily sessions
0
2
4
6
8
0
5
10

teacher trainer used a check list of 16 teaching techniques and simply marked the specific teaching techniques that the student instructor used. Throughout this study the teacher trainer provided the student instructor with any teaching material or data procurement device he requested. Criterion for completion of this study was similar to the others--two consecutive no-error days.

Results

Criterion of learning performance (two consecutive no-error days) was realized in six sessions. Figure 2 shows that the S's rate of correct responding progressively increased as his rate of errors steadily declined.

Of the 16 teaching techniques that were checked daily, the student teacher performed eight of them at all times; one occasionally, one in an altered form, and six, never (see Table I). At no time did the instructor convert drill time to seconds (he simply recorded the time to the nearest minute), record the pupil's errors and correct responses on the data sheet, calculate the correct rate, compute the error rate, record these rates on the data sheet, or graph them. The student instructor altered the method of tallying his pupil's responses by merely counting frequency of errors, rather than by marking the specific words missed. It was further noted that three times at the end of the drill session, the instructor failed to scramble the 20 words before testing the pupil.

Experiment VI

The intent of this study was to reteach and review the teaching procedures. The classroom teacher now acted as the pupil, as in Experiment III, in order to prompt the student instructor and to reinforce his correct use of teaching procedures. Again the criterion of two consecutive no-error sessions was used. During the first session, lasting for 30 minutes, the student instructor correctly performed 12 of 16 procedures, failing to calculate, record, or graph correct and error rates. During trials 2 and 3, which lasted for 20 and 14 minutes respectively, the student performed all the techniques correctly. Following this training, the student instructor was absent for three days and was unable to begin his teaching. One more training session was therefore administered to determine whether he would still follow the prescribed techniques. His performance was perfect during this final session, which lasted for 26 minutes. The total time of the 5 sessions of Experiment VI was 90 minutes.

Table I

STUDENT TEACHER

CHECK LIST OF TEACHING PROCEDURES

	SESSIONS						
	1	2	3	4	5	6	7
Mark session number	x	x	x	x	x	x	
Start watch	x	x	x	x	x	x	
Present cards	x	x	x	x	x	x	
Confirmation	x	x	x	x	x	x	
Stop watch	x	x	x	x	x	x	
Convert time to seconds							
Record time	x	x	x	x	x	x	
Scramble words	x	x	x				
Tally responses	*	*	*	*	*	*	*
Record on data sheet							
Compute earned points	x	x	x	x	x	x	x
Send S back to seat	x	x	x	x	x	x	x
Calculate correct rate/minute							
Calculate error rate/minute							
Record on data sheet							
Graph correct and error rate							

x correct use of procedure

* altered use of procedure

Experiment VII

Procedures

The intent now was to determine if these procedures would generalize when the same student instructor worked in the area of reading with a different pupil. The student instructor was given a set of 20 words to teach his new pupil. He was told that he would have 10 minutes for each drill session and that he could use the teaching procedures of his choice. He was also advised that he could terminate the study when he believed his pupil had learned the words.

Again two dependent variables were maintained daily, one pertaining to product and one to process. Each day the pupil's rates of correct and error responding were noted, while daily the teacher trainer noted on her check sheet any of the 16 techniques used by the student instructor.

Results

The student instructor continued the project for 6 sessions or until 3 consecutive no-error days were recorded (one more than previously required for criterion performance). During the first posttest the pupil mispronounced 9 words, obtaining a correct rate of one answer per minute. His correct rate during the final session, however, as 7.9 answers per minute (see Figure 2).

Throughout the drill and posttest, the classroom teacher recorded the student instructor's use of the 16 teaching procedures. All 16 procedures were used during all 7 sessions, without any reminders from the classroom teacher. In certain instances, however, some procedures were altered by the student instructor. For example, when confirming his pupil's correct answers, the student instructor at times said "fair," "good," and "beautiful" rather than merely "good." At other times, when his pupil could not pronounce a word, he would give descriptive cues rather than merely tell the pupil the unknown word. For example, when the word "patient" was not pronounced the instructor prompted, "What is in a hospital?"

Discussion

When the results of the seven experiments were summarized, it was noted that the student instructor taught 90 skills--80 reading words and 10 multiplication problems. He taught these skills in 557 minutes and was assisted by the teacher trainer for 752 minutes. As indicated by Table 2, the

Table 2

STUDENT TEACHER

SUMMARY

STUDY NO.	SKILLS TAUGHT	MINUTES	TEACHER TRAINING MINUTES
1	20 words	160	23
2	20 words	54	7
3			289
4	20 words	124	124
5	10 math problems	105	105
6			90
7	20 words	114	114
TOTAL	90	557	752

STUDENT TEACHING RATE: .16 skills per minute
1 skill per 6 minutes

TEACHER TRAINING RATE: .11 skills per minute
1 skill per 9 minutes

student's instructing rate was .16 skills per minute or 1 skill per 6 minutes. The teacher training rate was .11 skills per minute or 1 skill per 9 minutes.

The data from this series of experiments has provided some initial evidence that if the product of teaching is measured in terms of pupil performance and the procedures of teaching are first defined and then taught, a "teacher" can be prepared in a relatively short period of time.

One could argue, however, that the 16 procedures programed for this teacher do not encompass the scope and breadth of a true master teacher. Assuredly, other skills remain to be considered before a prospective instructor is presented with his teaching credentials. Certainly he must be able to assess pupil performance, specify educational objectives for his pupils, and finally, design remediation or therapeutic methods to obtain these objectives. Beyond these skills, it might be posited that an effective teacher is one who has some knowledge of current and past educational philosophies, understands and appreciates contemporary music and art, and participates in civic affairs. If these latter skills are deemed necessary for the teacher, it would be reasonable to program their acquisition in much the same way that the procedural skills of this study were arranged.

The implications of such a teacher training plan appear obvious. For if a teacher were evaluated on the basis of pupil performance (plus, perhaps, some additional personal behaviors) and the trainer's steps leading toward this realization were clearly described and sequenced, individualized teacher training, as well as individualized teaching, would be possible. Prospective teachers could advance through such a sequence at their own rate and would receive their credentials only when they had demonstrated their competency to modify pupil behavior effectively and efficiently.

Study IV: Effects of Manipulating an Antecedent Event on Mathematics Response Rate

Recent investigations have reported that certain academic behaviors have been increased as a result of manipulating specific subsequent events--those happenings that follow academic performance. McKenzie, et al. (1968)

studied the effects of grades and allowances contingent on the academic behaviors of a class of children with learning disabilities. They reported that when such subsequent events were used within a response-contingent framework, the rate of task completion was positively affected. Lovitt, et al. (1968) sought to determine the functional relationship between spelling accuracy and contingent leisure time. The results of this study, conducted with a class of fourth-graders, revealed that spelling performance improved when leisure time was systematically employed as a subsequent event.

Currently, many teachers attend to the arrangement of subsequent events in their efforts to alter certain academic behaviors of children. Although the scheduling of subsequent happenings is often random and unsystematic, classroom teachers often withhold such prized activities as recess or free time until the children have completed specified academic tasks.

Other teachers, however, particularly those attempting to encourage more accurate academic performance, manipulate antecedent events--those happenings prior to the pupils' responding. Many of these antecedent are in the form of instructions such as "Be more careful," or "Don't make so many mistakes." Another antecedent alteration, sometimes used by teachers attempting to effect more accurate performance, is to request children to use mnemonic devices. Children are given such aids as the "alphabet song" and the "30 days hath September" poem to recite before they write the letters of the alphabet or the number of days in each month.

Many of these antecedent alterations, such as instructions or memory aids, could positively affect academic performance. However, evaluations of their function are generally casual; rarely have these common classroom teaching aids been subjected to an experimental analysis.

The currect study was prompted by a student who exhibited extreme variability in responding to mathematics materials. His teacher had noted that at times he could answer certain problems correctly, while at other times he would err when presented with the same material. The teacher had further observed that occasionally, when the subject verbalized the problem before writing the answer, the probability of correct responding was increased. The purpose, then, of this investigation was to assess experimentally the function of a popular remediation suggestion of teachers and parents, that children "think before they do." More specifically, this study was designed to determine the effects of simply writing the answers to mathematics problems versus the effects of verbalizing the problems prior to making a written response.

Subject and Conditions

The subject was an 11-year-old boy attending a class for children with behavior disorders at the Experimental Education Unit of the University of Washington.

Procedures

This investigation comprised three individual experiments. For each session the subject was placed in a small room furnished with a table, a chair, and an additional seat. Each day the experimenter provided him with a pencil and a sheet of math problems. Sessions were timed by the experimenter, who sat next to the subject with a stop watch.

During the 26 sessions of the first study, the experimenter daily presented the subject with a set of 20 mathematics problems of the class $\square - 2 = 6$ where the minuend and subtrahend $\leq$ 10. A pool of 121 problems of this class were available. From this collection, 26 separate sets were constructed, composed of 20 randomly selected problems. The basic considerations in making up each set were that all the problems on a single set would be different and that identical problems would not appear in 2 consecutive sets. Throughout this experiment each problem reappeared from 4 to 7 times.

During the first phase, the experimenter told the subject to write the answers. When the subject finished a page, the experimenter made a comment such as "Thank you, that's fine," calculated the total time of the session, then returned the pupil to the classroom. The examiner then totaled the correct and error responses and converted these figures to rate per minute by dividing each total by the time of the session.

In the second phase, the manipulation phase, the subject was told to say each problem aloud before writing the answer. He was first given a practice problem, with the instructions, "Read the problems aloud like this: 'Some number minus two equals six.' Then write the answer. Say this problem first to practice: $\square - 2 = 6$." After starting the page, neither the subject's verbalizations nor his answers were corrected; he was, however, reminded to say a problem aloud if he failed to do so. After the third day the practice problem was no longer required, since no examiner prompts were necessary during the third session.

During the final portion of the experiment, phase 3, the subject was told that he no longer needed to say the problem aloud before writing the answers. If he did verbalize the problem, the experimenter reminded him to merely write the answer, without reciting the problem.

Following the completion of the first study, Experiment II was conducted. This investigation, consisting of 30 sessions with 20 problems presented per session, was designed to replicate the previous experiment by again assessing the subject's performance as a function of either his reciting or not reciting the problem prior to his written response. During Experiment II a collection of 60 problems was available of the class $\square - 20 = 40$, where the minuend and subtrahend were of the tens class from 0-100. From this pool, 6 sets of 20 problems each were constructed. All the problems within the sets were different, and although each problem was included more than once in the series, identical problems did not appear in 2 consecutive sets. These sets were then presented on consecutive days, beginning with the first set and continuing through the sixth, then beginning again with the first. Throughout Experiment II each problem was thus presented 8 to 12 times.

In phase 1, the subject was told to simply write the answers to the problems. During phase 2, he was instructed to verbalize "Some number minus twenty equals forty," prior to making a written response. Finally, in phase 3, the subject was instructed that he should not verbalize the problem prior to writing his response. All other procedures remained identical to Experiment I.

Experiment III consisted of 32 sessions. Throughout this experiment 10 problems, of a class more complex than those of the preceding two experiments, were offered at each session. Seven sets, each containing 10 problems, were constructed from a pool of 70 items of the class $4 - 3 = 9 - \square$. These sets were then scheduled on consecutive days, beginning with the first and continuing through the seventh, then beginning again with the first. The sets were arranged so that the problems within and between each set were different. During the 32-day experiment, then, the same problems reappeared each seventh session or about 4 times throughout the experiment. All procedures remained identical to those in Experiments I and II.

Results

Experiment I

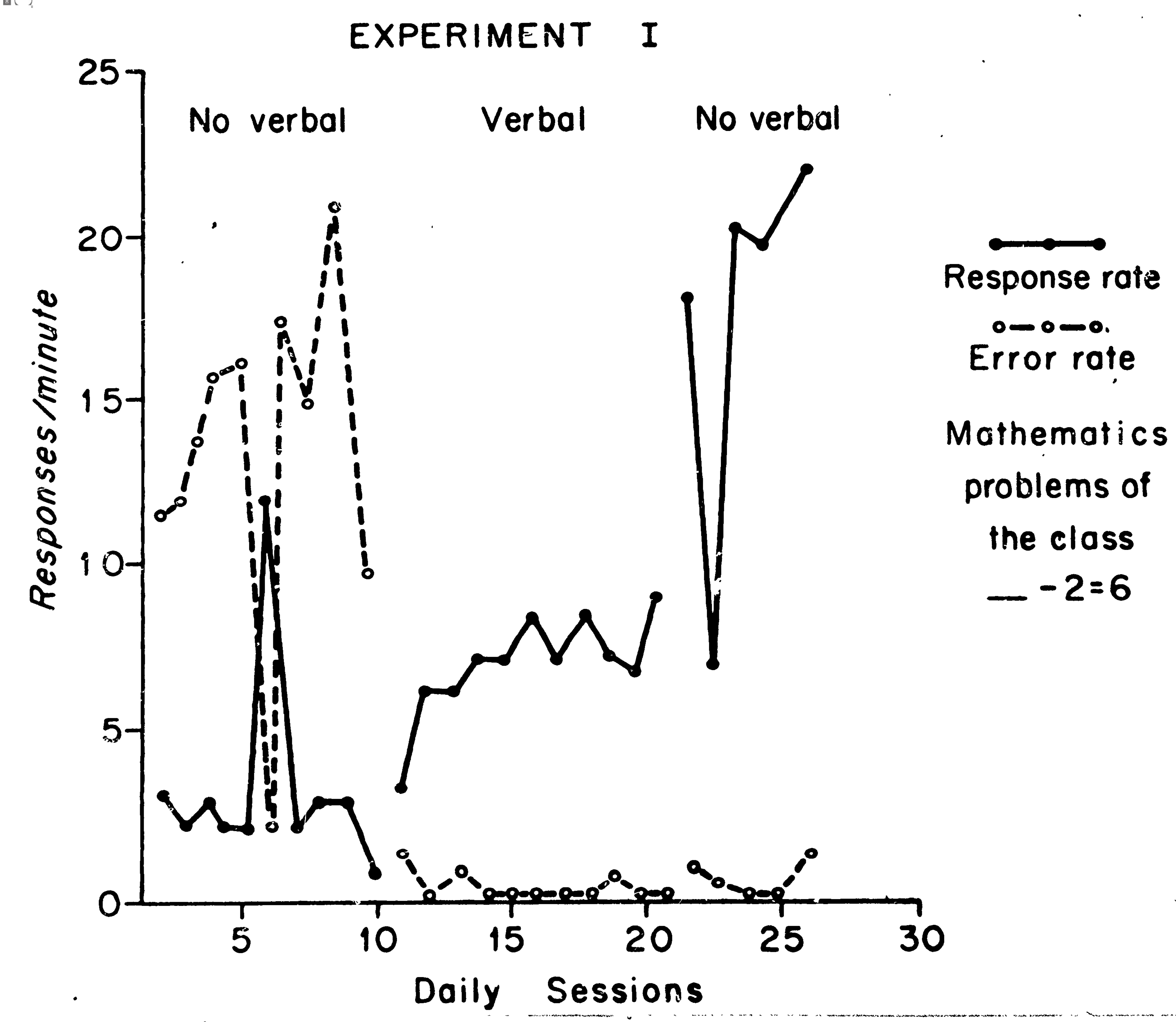
In the initial phase, where no verbalization was required of the subject, his correct rate ranged from .6 to 12.0 with a mean of 3.0 responses per minute (Figure I). His mean error rate at this time was 13.4 with scores ranging from 1.8 to 21.0. When the subject was required to verbalize the problem prior to responding in phase 2, his correct rate increased to a mean of 7.1 with a range of 3.0 to 9.0. At the same time, his error rate decreased to a mean of .2 with a range from 0 to 1.2. In the final phase, although verbalization was no longer required, the subject's correct rate continued to increase, reaching a mean of 17.4 and ranging from 6.6 to 22.2. His error rate during this latter phase continued to be low (mean=.6, range=0-1.2).

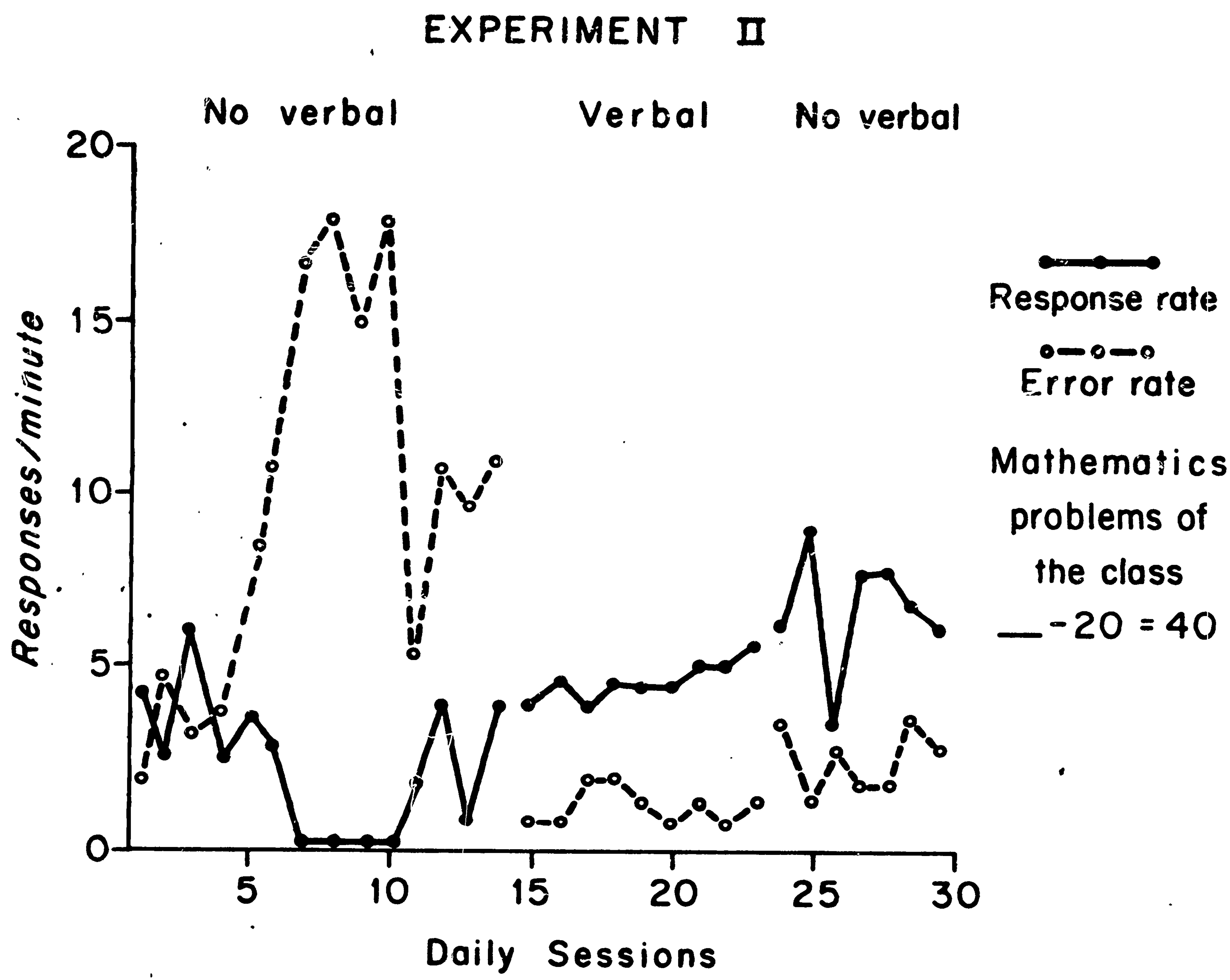
Experiment II

The results of this study paralleled those of Experiment I. During the manipulation phase the subject's correct rate rose and his error rate decreased from the nonverbal initial phase. The subject's correct rate increased still further, as in Experiment I, in the final nonverbal period. Mean correct rates were 2.2 for phase I (range=0-6.0); phase 2, 4.3 (range=3.6-4.8); and phase 3, 6.7 (range=3.6-90.0). As indicated by Figure 2, the mean error rates in the second experiment decreased from 9.8 (range-1.8-18.0) in phase I, to 1.1 (range=.6-1.8) in phase 2, the manipulation period. In the final phase, phase 3, the mean error rate rose slightly to 2.4 with a range from 1.2 to 3.6.

Experiment III

In the final study, the subject again increased his rate of correct responding during and after the manipulation phase. The subject's mean rate of correct responding in the initial phase was 1.4 (range=0-3.0). In phase 2, this correct rate increased to an average of 3.9 with a range from 2.6-5.4; and in phase 3, it increased still further to a mear of 8.7 ranging from 5.6 to 11.1. Error rates decreased between phases I and 2,


EXPERIMENT I
No verbal
Verbal
No verbal
Responses /minute
Daily Sessions
Response rate
Error rate
Mathematics problems of the class
-2=6


EXPERIMENT II
No verbal
Verbal
No verbal
Responses/minute
20
15
10
5
0
5
10
15
20
25
30
Daily Sessions
Response rate
Error rate
Mathematics problems of the class
-20 = 40

declining from 11.1 (range=4.8-27.0) to .2 (range=0-.6). The subject's error rate in phase 3 increased slightly to a mean of .8 (range=0-5.6).

Discussion

Throughout the three experiments of this study, the subject's pattern of responding was relatively consistent. During the initial phase of each experiment, when the subject simply wrote the answer to the problem without verbally stating the problem, his error rate was much higher than his correct rate of response. During the second phase of each experiment, the verbalization period, his error rate dropped considerably, generally near zero, while his rate of correct responding increased. When the subject no longer verbalized the problem (phase 3 of each experiment), his rate of correct responding increased by a large amount while his rate of committing errors increased only slightly. Furthermore, the subject's variability was much lower during the verbal phase of each experiment. This may be noted by comparing the subject's range of correct and error rates during the second phase of each study with phases 1 and 3.

Since the same problems were used repeatedly within each experiment and since the pupil's rate of correct responding actually increased from the verbal second phase to the nonverbal third phase, it could be said that the effect of verbailization generalized. Generalization, however, was restricted to mathematics problems of the same class; when different problems were introduced during subsequent experiments, the student's accuracy was again low. This is evidenced by a comparison of the pupil's correct rate of responding in the final phase of Experiment I with the first phase of Experiment II, and the final phase of Experiment II with the initial phase of Experiment III.

Such cases, where generalization does not occur from one situation to another, pose problems for teachers. In such instances, the teacher must systematically arrange and develop each sequential step; for as Baer (1968) pointed out, generalization "should be programed, rather than expected or lamented."

The current study, which experimentally assessed the function of a popular remediation theory--that verbalizing a problem aids in solving it--revealed data supporting this theory. These results may be due to one of a number of factors. One possible explanation may be that the subject's increased accuracy when the problem was verbalized was a function of adding a second stimulus dimension to the mathematics problem--orally reciting the problem as well as seeing it. The theoretical basis for such a thesis, that learning proceeds at a more rapid rate under multi-sensory conditions, has historically been supported by numerous educators.

An alternate explanation for the subject's more accurate performance during and following periods of verbalization could be that he simply became more deliberate. A pupil, when required merely to write an answer to a problem, could regard the effort of making a pencil mark as extremely minimal. However, when required to read a problem aloud and then to write the solution, the increase in complexity could have functioned to make the pupil more deliberate.

Whether saying the problem before making a written answer served to increase the pupil's accuracy because of multi-sensory involvement, or whether the pupil became more deliberate as the process became more complex, are questions which these data do not presume to answer. What is important is that, in this study, when the variable--a verbal statement of the problem--was manipulated, it was found to be effective in improving arithmetic performance. Furthermore, this report demonstrates that the individual effects of various remediation theories and maximx might be determined through the use of experimental analysis procedures.

References

Baer, D. M., Wolf, M. M. and Risley, T. R. Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1968, 1, 91-97.

Ferster, C. B. and Skinner, B. F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.

Haring, N. G. and Kunzelmann, H. P. The finer focus of therapeutic behavioral management. Educational therapy, Vol. 1. Seattle: Special Child Publication, 1966, 225-251.

Haring, N. G. and Lovitt, T. C. Operant methodology and educational technology in special education. In N. G. Haring and R. S. Schiefelbusch (Eds.) Methods in special education. New York: McGraw-Hill, 1967, 12-48.

Lovitt, T. C. and Curtiss, K. A. Choice as an independent variable. Unpublished manuscript, University of Washington, 1968.

Lovitt, T. C., Guppy, T. E., and Blattner, J. E. The use of a free-time contingency with fourth graders to increase spelling accuracy. Behavior Research and Therapy, in press.

McKenzie, H. S., Clark, M. Wolf, M. M., Kothera, R. and Benson, C. Behavior modification of children with learning disabilities using grades as tokens and allowances as back up reinforcers. Exceptional Children, 1968, 34, 745-753.

Morse, W. H. Intermittend reinforcement. In W. K. Honig (Ed.) Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966, 52-108.

Singer Mathematics Program. Sets and Numbers. Patrick Suppes, Director. New York: L. W. Singer, 1965.

Sullivan Associates Program, Programmed Reading. Cynthia Buchanan, Director. New York: Webster Division, McGraw-Hill, 1963.

CHAPTER XIII A TACTIC FOR ADMINISTRATION OF SPECIAL CLASSES

H. P. Kunzelmann

Introduction

The school administrator daily faces problems related to teacher personnel, primarily because incidents such as teacher illness, demands on teacher time for other assignments, parental concerns, and disciplinary reasons all draw the teacher from the classroom for varying periods of time. Whether a teacher is gone from the class for 15 to 20 minutes or the entire day, or is absent because of illness or emergency leave, the administrator must be aware of the actual effects that the teacher's absence has on the class members. The exposure of various contingency systems, management techniques, and teaching styles by different substitutes for short periods of time might have effects on children that are unknown at present due to the lack of continuous assessment data. The administrator, or anyone concerned with changes in classroom management, would be justified in insisting that more detailed lesson planning or teaching procedures be pre-scheduled in any classroom, whether or not the daily performances of the children are adversely affected by the change of teachers.

A second area with problems relevant to an unplanned teacher exchange lies in the attempt to identify testable teaching procedures. First, the program plan for subject matter presentation needs some consistency within the bounds of the classroom. Subject matter presentation may be handled by pupils, teachers, peers, machines. But, in all cases, the consistency with which the program is presented, the style (verbal, visual, or both) used for presentation, when held consistent, may not produce variability in pupil performance.

The third area of testable teaching procedures is the contingency and consequence system that the teacher utilizes in the classroom. Teaching acts, such as correcting children's work, giving numerical values to correct and error work, recording the scores that the children make, are examples of contingency and consequence systems. Are the results given to the child consistent within each subject area? Is feedback immediate in some areas and not in other subjects? These contingency and consequence systems can be tested, provided that the total teaching procedure is outlined prior to the teacher's confrontation with the children and insistence upon given performance rates.

A fourth problem area that the administrator faces is the effect of school calendar events on the performance of pupils. The vacation periods that systematically interrupt continuous teaching for social reasons probably produce one variable for change. Another is the pupil absenteeism for valid or invalid reasons. The child is not performing on the skills designated by the teacher, not following a systematic format, and consequently is out of step with the general class membership. The issue here is not one of changing calendar events but in having a technique for accountability.

The issues above are everyday problems without everyday answers. The usual decision is to ignore the issues because of a lack of procedures for assessment. The effects on pupils when teachers are different for the same performance area have not been explored. At best, we find researchers concerned with such issues as teacher personality and model teaching.

Common Issues Relevant to Pupil Performance

The development of rate measures for continuous classroom measurement was reported by Nolen, Kunzelmann, and Haring in 1967. The rate measures described therein were found to be extremely sensitive indicators of pupil performance on academic material in an intermediate classroom. The recording of classroom behavior accomplished by the teacher or the pupil was described as a continuous assessment procedure. The study was conducted in an intermediate-age-level class of approximately nine pupils. The recording system for the study consisted of two movement cycle counts and one time measure to record pupil performances. The movement cycle counts were separated into areas of correct and error counts. The study used six control variables managed by the pupil or the teaching, occurring either before or after a child's performance, all of which could be independently managed. The recording system described pinpointed the actual contingencies used by either pupil or teacher on the six controls. The six controls were: 1) teacher-initiated contact; 2) pupil initiated contact; 3) subject presentation by either pupil or teacher; 4) points based on teacher or pupil correction given by either pupil or teacher for contingency controls; 5) content of the subsequent events or the environmental occurrences after the child behaved; 6) the time at which the subsequent events could be used by the pupil. These six independent control measures in the classroom developed into an extremely refined system where, at the present time, some 192 teaching styles can be used for complex contingency management for any classroom of children.

Combining specified elements that may show function for the control of classroom performance is precisely what the teachers, Mr. William Hulten and Mrs. Rita Ballard, secondary class; Miss Karen Curtiss and Mrs. Judy Burgess, intermediate class; and Mrs. Marilyn Cohen and Mrs. Patricia Fosnick, primary class; managed in the study reported. Although each classroom team had few teacher absences, when they did, it was not a potential assessment period because most of the time the assistant teachers or head teachers would follow the previously outlined teaching plan. Testing each lesson plan was then a matter of testing the contingency management system. If pupils were performing movement cycles under specified classroom conditions, it was relevant to know if the pupils' individual rates were accelerated, maintained, or decelerated.* The data

* *Maintenance was, at the time of this study, considered excellent for it was assumed, wrongly, that progression through materials, i.e., Book 1 to Book 2, was more difficult; subsequently, maintenance was supposedly ideal while the child "went through" presumably more difficult materials.*

included here depicts maintenance. The teachers were aware of their own achievements in managing a classroom, as well as the progress of the pupils.

Procedures for the Investigation of the Effects of Exchanging Teachers within Classes

The 22 pupils involved in the study ranged in age from 7 to 17 years, and had been enrolled in the classroom programs for period of 1 month to 2 years. The pupils performed generally in subject areas of academic request by the teacher at rates of approximately .5 to 10 movements per minute. The classrooms were divided into two areas, low and high strength. This division enabled the teacher to conduct activities around low-strength or low-performance areas while keeping the high-strength or high-performance areas contingent upon a child's low-strength production. Each classroom kept continuous performance data for about 2 years with the last year of the data-keeping recorded as movements per minute.

A more refined description of the content of the classroom would require information about each child's program for, in all the classrooms, the children were individually programmed. Such detail will not be included except to say that the children in each of the classrooms were at varied levels. A <u>before</u> phase, prior to general change, can best be described by the type of data gathered in the class, the antecedent conditions that the teacher utilized for the presentation of material, the arrangements that the teacher set up to manage the pupil's subsequent events, and the actual content and system of subsequent events in the classroom. The before phase can be further exemplified by the notation in Figures 1-6 where the before phase would be the first 2 or 4 weeks prior to any systematic changes or planned changes that were scheduled to occur in the class.

The type of data gathered had a topography of both oral and written answers to arithmetic problems and oral or written answers to reading problems. The children arrived at approximately 9:00 a.m. every day and remained until 2:00 p.m.--five hours or 300 minutes of potential instruction time. In all classrooms, math and reading were usually conducted in the morning hours between 9:00 a.m. and noon. The study concentrated on the 180 minutes available between 9 a.m. and 12 noon daily so as to report on the two basic subjects, math and reading.

The math curriculum consisted of the Suppes material, Sets and Numbers, ranging between Book IA and 5B. The reading materials were basically the Sullivan Reading Series where in the program format the child would have to make both oral and written responses. The antecedent conditions were constant throughout the school day. The teacher presented the child with the work at the appropriate page, advising him to commence his exercises either in math or in reading. Specific antecedents were handled by materials presented by teachers. In the intermediate class, some children were allowed to select the materials they wanted; the time of the day was held constant after the first 2 to 4 weeks. The classroom arrangements were managed so as to ensure that each child would be handled in a consistent fashion after he had committed himself to a given number of problems.

As an example, if a child were to finish one page or approximately fifteen written responses in a Sullivan Reading Program, he might be told to raise his hand to signal a teacher. This variable of teacher specification of pupil-initiated contact would then enable the teacher to correct the fifteen answers that the child had made, determining the number of correct and incorrect responses and then giving the child some point value for the number of correct answers. The child might be placed on a 10 to 1 schedule with 10 correct responses redeemable for one minute of free time. If the teacher were to find fifteen correct problems on a child's work page, the child would begin 1 1/2 minutes free time. Such classroom arrangements for management and contingencies were maintained in all the classrooms.

In all the classrooms the subsequent events consisted of free time activities usually chosen by the children. The children generally were allowed to have an array of possible events that they could manage for themselves, such as science kits, games, picture books to read, records for listening, and radios to play. All the available activities that children would either request or the teacher would test for function would be in the high-strength area of the classrooms. The subsequent event menu consisted of activities that the child could indulge in while utilizing earned minutes of free time. In most classrooms the children were allowed to take their free time when they requested it. However, if the teacher had arranged the subsequent events to occur at a particular time of the day for a more organized activity, a fixed system of subsequent events evolved.

Throughout phase I, each teacher was able to hold to a continuous plan for assessing the effects of the program on the children's performance rates. The antecedent conditions in the classroom--the time of day, the amount presented, the types of materials, the consistency with which the children were presented material--all were fixed variables. The arrangement system, the contact initiation, the number of points for correct responses, as well as the subsequent events selection plan were all managed in a consistent fashion and held constant throughout the two- to four-week periods prior to the announcement of a teacher exchange.

The additional variable placed on the three classrooms was at a teacher's meeting on April 19, 1968 where it was announced that a teacher exchange week would occur two weeks from the following Monday. This was the first formal notice the teachers received concerning the possibility of their class being assigned to a different teacher for a period of one week while they, in turn, would take over the class of another teacher. Each teacher would have the option of sending an assistant teacher to the classroom for approximately a three-day orientation period during which the regular teacher would outline the precise management procedures in the classroom to the exchange assistant teacher. Of deepest concern to all the teachers was the procedure for social contingencies that had been established in the room to manage non-academic behavior. For example, if a child talked out in the classroom, what was to happen? Was the child to be ignored? Were the other students to be complimented on not attending to the child or scolded for attending to the child? Was the teacher to issue the child an immediate subsequent such as a face mask? Such questions arose at the teacher meeting when the change was discussed. Each of the teachers was able to identify quickly all the contingency plans and subsequently to pass on instructions concerning them to the assistant teacher who would be involved in the class during the week of the exchange.

The procedures used for the exchange follow: The teachers, by a matching system, decided in which classroom they would assume responsibility and what teacher would take their own class. The amount of information exchanged by the teachers was to be as thorough as possible. It was agreed that the secondary class teachers would draw the primary class; the primary class teachers, the intermediate class; and the intermediate class teachers, the secondary class. This arrangement was a random decision made during the teacher meeting where the exchange was discussed. It was immediately noticeable that the amount of information to be shared would be much more complete than the teacher's lesson plan within the classroom where each child's contingency management system provided for individualization. The agreement about the amount of information to be exchanged was nebulous at the time and would have to evolve during the first two weeks prior to the actual exchange. The dates of the exchange were governed mainly by the calendar, the same problem mentioned in the introduction to this paper. The teachers would routinely have to refer to the classroom's general structure of operation on the ten-week university quarter. This happened to be the spring quarter and the exchange would have to take place one week before the end of spring quarter so that we could observe what occurred when the teachers returned to their regular classes.

A few general comments as to the exchange: There seemed to be an increase of verbal behavior as the time of the exchange week approached. There were correction problems between teachers about whether certain problems would be marked right or wrong according to the teacher's criteria. It was hoped that through the consistent use of the Sullivan Reading Series and the Suppes Modern Math Series the teachers, who had worked on various levels, would have a consistent pattern for identifying correct and error movement cycles by the pupils. This was not totally the case. However, the data does indicate, as will be pointed out in the results section of this paper, that there could not have been a tremendous amount of difference in the actual grading system. The concerns expressed by the teachers reflect in the pupil's data, as will be shown later. The classroom teachers became extremely "clean" in their management procedures, specifying each contingency arrangement to ensure that each child was handled consistently. The information about each child's programmed antecedent condition and subsequent activities was passed on exactly as the teachers had managed them in the before phase. The exchange week brought about little verbal teacher interaction, the teachers being content to allow the performance rates to indicate the success or failure of the experiment.

The last phase of the procedure was the teachers' return to their regular assignments. This of course is a replication of the before phase described earlier in this section. Upon return to the classroom, the teachers noted effects which will be discussed in the results section. However, a description of the before phase should suffice to describe the contents of regular assignments in the last stage of the experimental sessions. A note about the teachers involved in the experiment: The three master teachers involved in the experiment had at least two years' experience as teachers for the Experimental Education Unit School program. They had been trained primarily at the University of Washington in Special Education Graduate Courses with subsequent training for their practicum or assistant teacher status at the Experimental Education Unit. During these two years, their experiences included managing children of all behavioral descriptions. Problems ranged from seizure management to bizarre verbal behavior of children in the classrooms.

Each classroom had at least one severe child so that management was a continual nebulous issue. The teachers' experience in contingency management was probably as refined as could be found. Each teacher at any moment was able to pinpoint the exact behavior of the child that would subsequently bring in an environmental event managed either by the teacher, the pupil himself, or a peer. All the teachers had refined training in the use of subsequent events as possible consequences or reinforcing events from the environment. During the two years, each had explored multiple arrangements for subsequent events in the classroom, detecting what the children liked to cash in their points for as well as the continuous management of the high-strength area in their classrooms. The teachers were well known to each other, not only through teacher meetings, but through social involvement over the two-year period. All the teachers had in common a lesson planning system that necessitated exact specifications of contingency management in the classroom. The teachers could best be described as persons who were highly concerned with each child's individual performances. Subsequently this concern was shared by their assistant teachers in each classroom.

When the announcement of the teacher exchange was made, the agreement criteria that teachers set for themselves during the discussion seemed to be reduced as the two-week period following the announcement of exchange progressed. Not only was there generally less variability in the childrens' performances, but this variability clearly seemed to be a function of some of the statements that the teachers continually made about the various ways in which the social management might falter. This is not to detract from the excellent quality of instruction administered by the teachers, but only to pinpoint clearly a possible effect of the announcement of the teacher exchange. This was unique in that a teacher exchange is not typically familiar to those of us who teach. Second, the relationships to the problems of substitute teaching, of absenteeism of the professional staff in a school, had not been systematically explored verbally. The teachers were confronted with a new factor in testing the effects of their particular management procedures. Their individualized, detailed lesson plans would be used by someone else.

The teachers were clearly aware of the magnitude of this study. Their particular lesson plan system was being challenged in that they certainly had to consider the questions, "Am I necessary for the classroom?" or, "If I can set up a classroom, must I be the person to operate it?" Each of the teachers knew also that a subtle question was being explored. In essence, the question was whether it is possible to let the children's performance specify whether or not the teacher who is conducting a very refined management system in the classroom can have the same precision management procedure continued when the teacher who evolved, developed and refined it is not present. Such questions are unique in that teachers as well as professional educators have been concerned with a nebulous notion of teacher personality and its relationship to effective teaching and only subsequently with the effectiveness of teaching in its relation to pupil performance. The relationships established between these three areas have not been clearly specified. The results of this study will clarify the lack of relationship.

Analysis of Results

The data indicate no basic changes for the classes. Figures 1-6 will be discussed in total by condition. The four conditions studied in detail are: 1) regular assignment 1, 2) announced exchange, 3) exchange week, and 4) regular assignment 2.

The analysis is detailed in three phases. First the difference in median changes between conditions according to the Lindsley-Mid-Middle Test. This test assesses the chance factors of changes. Second, the direction of the change as described by acceleration, increase in rate; or deceleration, a decrease in rate. Third, the variability of the data, its spread defined as range only, is discussed by visual inspection (less, more, and same) and summarized.

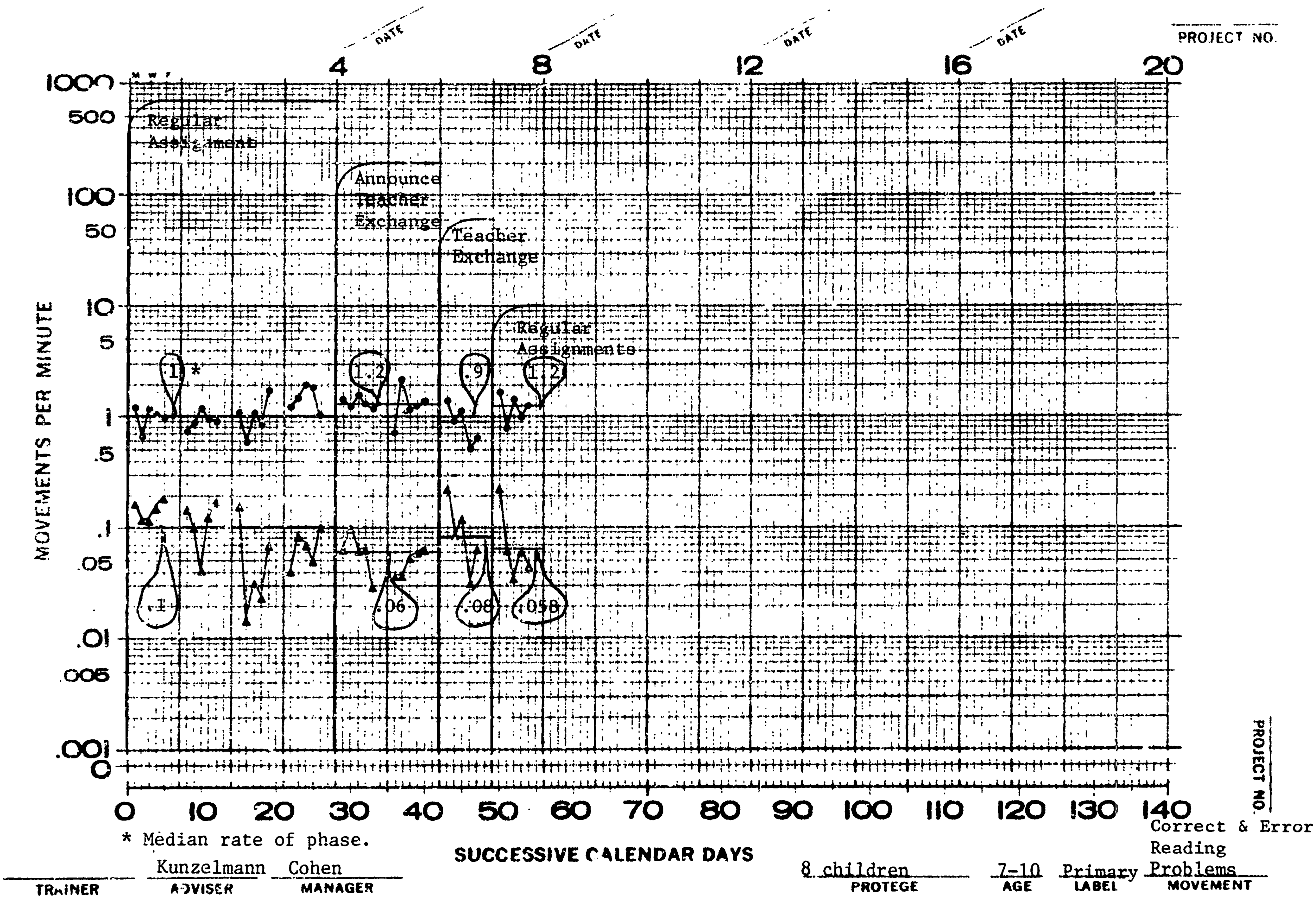
The primary class data (head teacher, Mrs. Marilyn Cohen) for reading (Figure 1) and math (Figure 2) reveal middle scores of 1.2 to .9 and 1.2 to .68 respectively for correct responses. Correct responses are always depicted by solid lines in the upper portion of the figure (errors are not analyzed).

Special note should be taken by the reader of Figure 2 conditions--exchange--and regular assignment #2 (last week of data) that the change would have only occurred by chance 4 times in 1,000 re-assessments. Also noteworthy during this initial analysis is that for both reading, Figure 1, and math, Figure 2, data direction was constant; i.e., accelerated--accelerated between regular assignment #1 and announcement of exchange. This directional cohesiveness occurred in all conditions and in both subjects in the primary room only.

Figures 3 and 4, data from the intermediate class (Miss Karen Curtiss, head teacher) reveal high middle scores of 1.2 to 1.0 for reading and 1.2 to .80 for math. The median rates of movements per minute reflect a small amount of variability except for the exchange week.

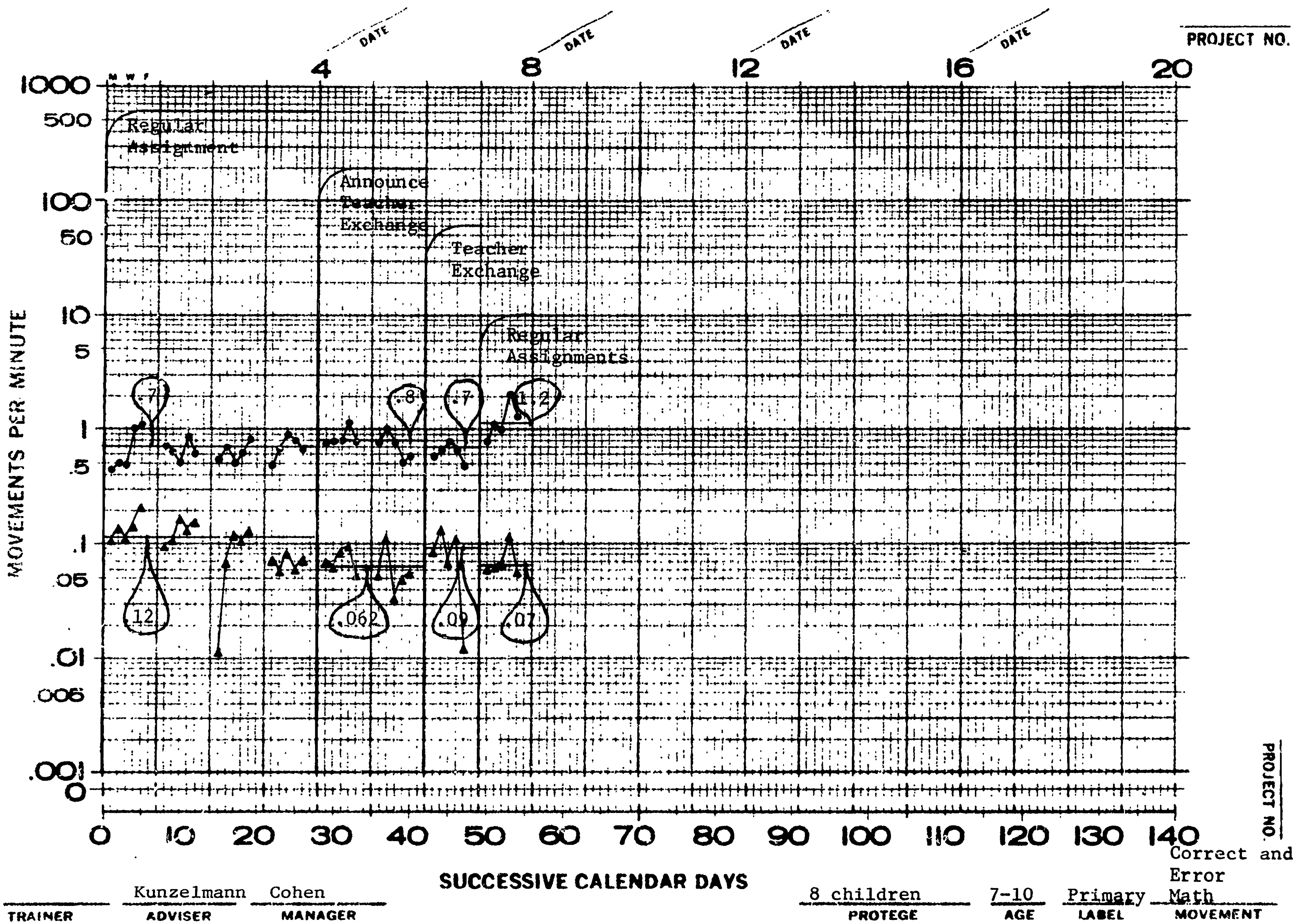
Of special interest for initial analysis is the constant middle scores per condition 1, 2 and 3, while in the regular assignment #2, the

Figure 1



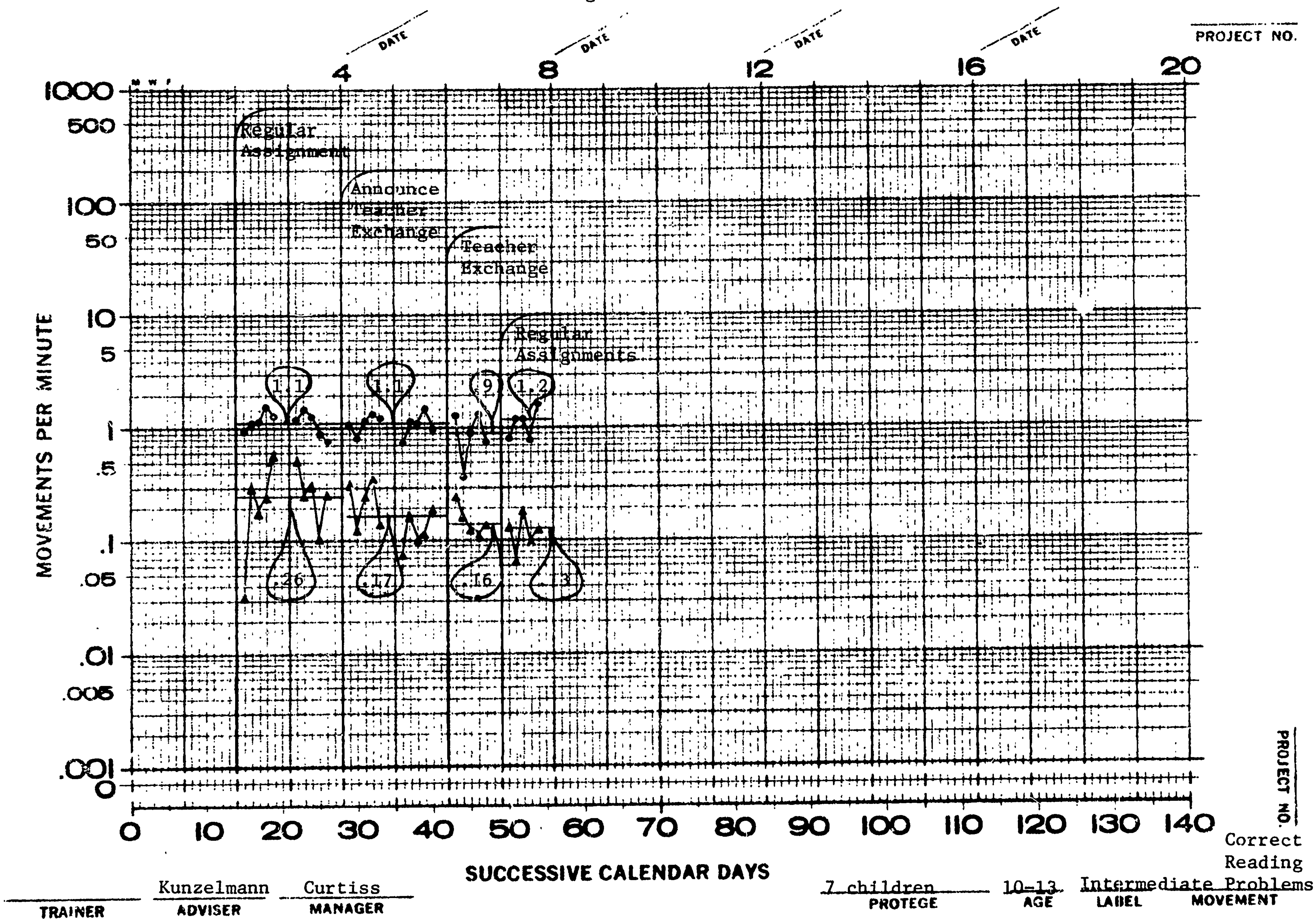
DATE
DATE
DATE
DATE
PROJECT NO.
4
8
12
16
20
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
M W F
Regular Assignment
Announce Teacher Exchange
Teacher Exchange
Regular Assignments
1 *
1.2
.9
1.2
.1
.06
.08
.058
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
* Median rate of phase.
SUCCESSIVE CALENDAR DAYS
TRAINER
Kunzelmann
ADVISER
Cohen
MANAGER
8 children
PROTEGE
7-10
AGE
Primary
LABEL
Correct & Error Reading Problems
MOVEMENT
PROJECT NO.

Figure 2


PROJECT NO.
DATE
DATE
DATE
DATE
4
8
12
16
20
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
Regular Assignment
Announce Teacher Exchange
Teacher Exchange
Regular Assignments
.7
.8
.7
1.2
.12
.062
.09
.07
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
PROJECT NO.
Correct and Error Math
Kunzelmann
Cohen
8 children
7-10
Primary
TRAINER
ADVISER
MANAGER
PROTEGE
AGE
LABEL
MOVEMENT

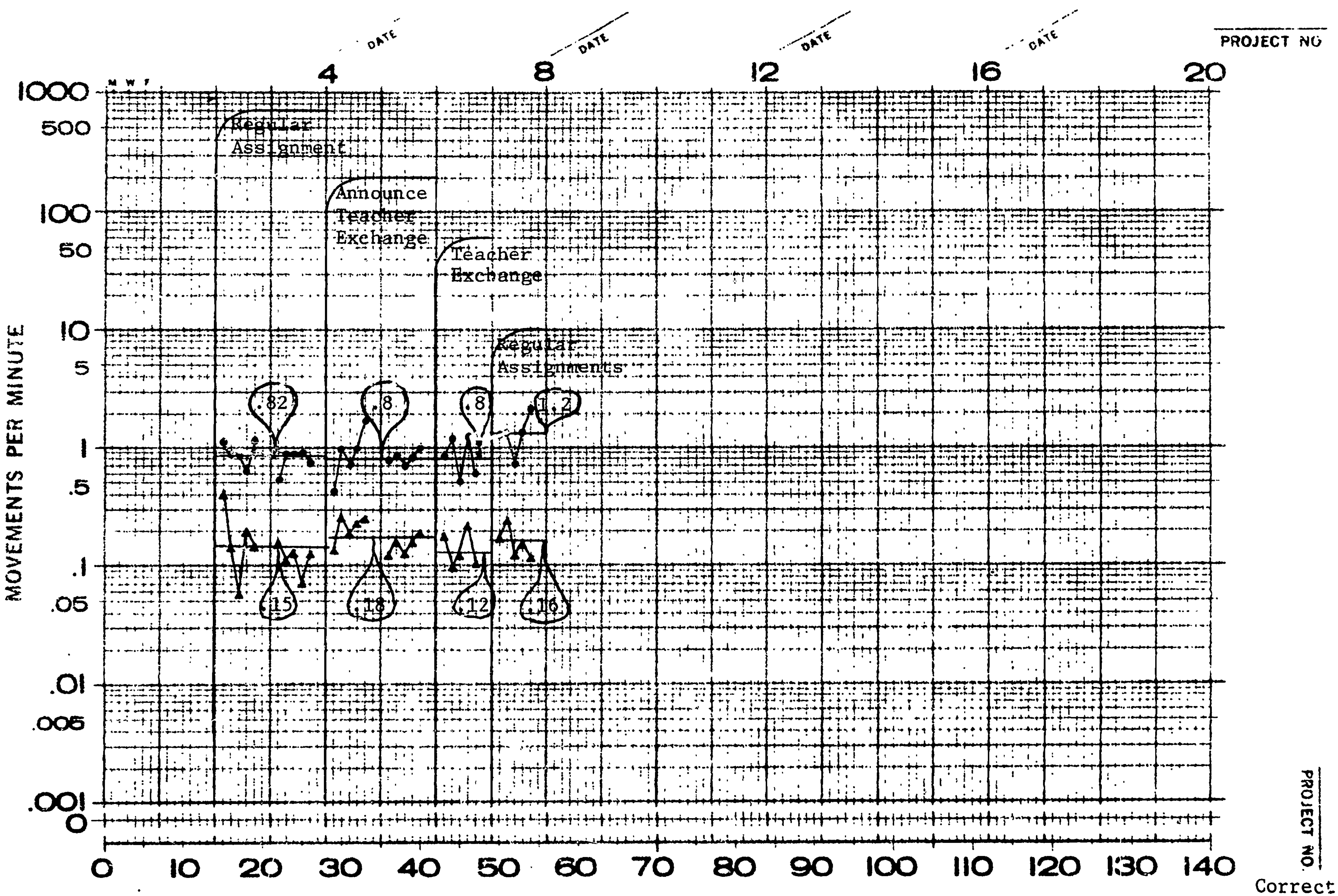
316

Figure 3



PROJECT NO.
DATE
DATE
DATE
DATE
4
8
12
16
20
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
Regular Assignment
Announce Teacher Exchange
Teacher Exchange
Regular Assignments
1.1
1.1
.9
1.2
.26
.17
.16
.3
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
TRAINER
Kunzelmann
ADVISER
Curtiss
MANAGER
7 children
PROTEGE
10-13
AGE
Intermediate
LABEL
Correct Reading Problems
MOVEMENT
PROJECT NO.

Figure 4



TRAINER: ______ ADVISER: Kunzelmann MANAGER: Curtiss PROTEGE: 7 children AGE: 10-13 LABEL: Intermediate MOVEMENT: Correct and Error Math last week reported, a marked change in math was noticeable when the regular teachers again took over the class. Also, it should be noted that p-values for announce and exchange week and exchange and regular assignment #2 were .08 and .05 in that order. More will be said of these differences later.

Figures 5 and 6, data from the secondary class (Mr. Bill Hulten, head teacher) reveal a low of .8 to a high of 1.1 in reading and, in math, a low of .6 to a high of 1.4. The two lowest middle scores occurred in the regular assignment #2 condition.

Special notation should be made that the data direction in the last week of reported data had in a marked deceleration pattern but would have occurred by chance in 10 out of 100 re-assessments.

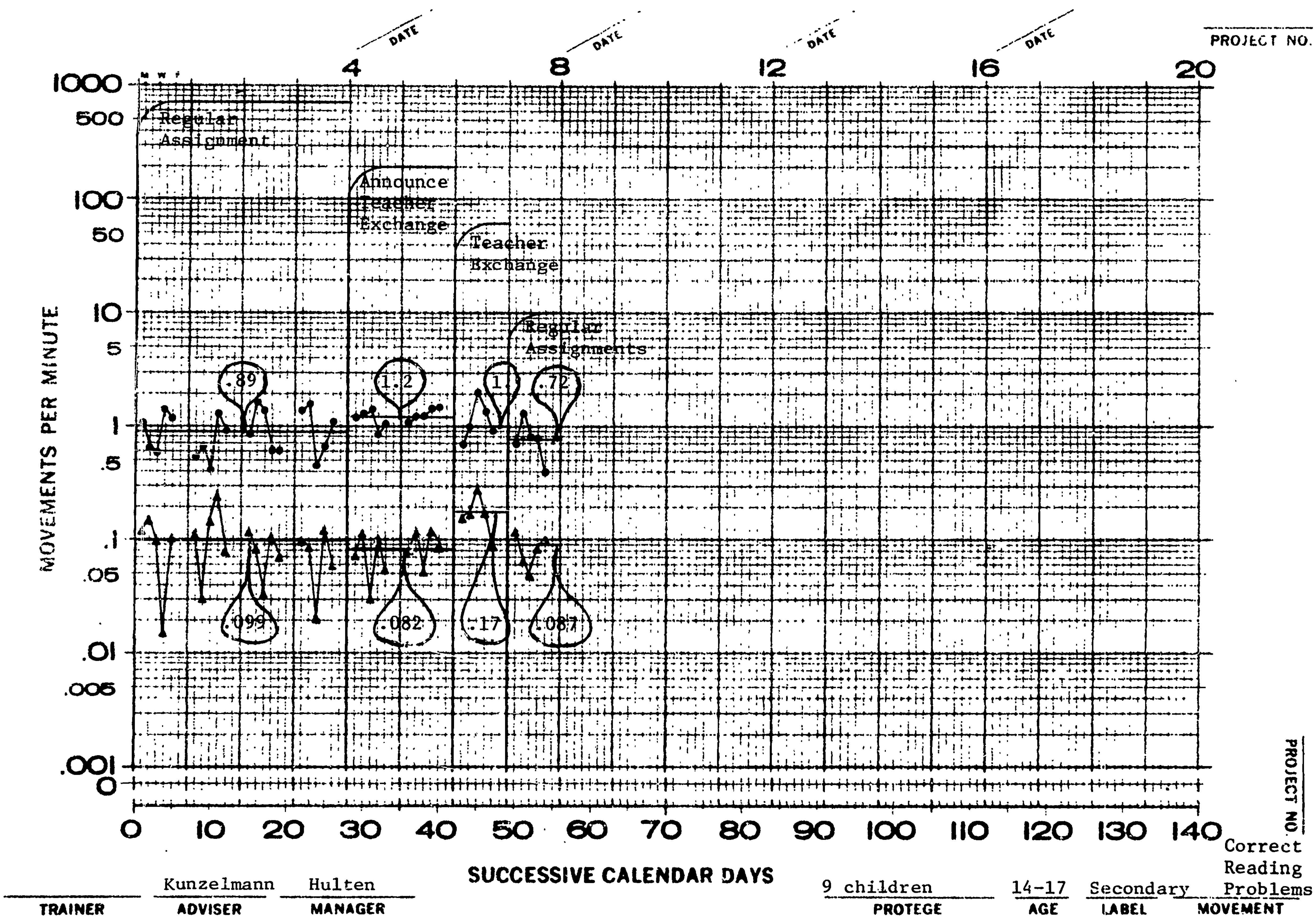
Figure 7, 8, 9 and 10 illustrate the rate ranges of all classes, correct and error in math and reading. Figure 7 shows that for the pre-announced exchange of 2 to 4 weeks, the range was .22 to 3.1 correct math problems per minute. The range that held 50 per cent of the rates was .6 to 1.1 problems per minute. During the announced exchange phase of two weeks the correct math problem rate ranged from .6 to 1.9 while 50 per cent of the plots fell between .8 and 1.1 per minute. The difference in variability of the plots indicates a narrowing effect or more control of the data during the announced exchange phase. Visual inspection clearly shows that control weakened during the next two phases of exchange and regular teacher assignment.

Without burdening the reader with more numbers, visual inspection of Figures 8, 9, and 10 shows that the same control takes place during the announced exchange phase for errors in math and correct-errors in reading.

To summarize the data, Tables I, II and III are provided. Table I shows the middle score for the classes of 22 children. The table reveals that both in reading and math the children emitted about one movement per minute. The reading rates are higher than the math rates in the primary and intermediate classes, while rates between subjects vary in the secondary class.

Table II lists the summary of the p-values between conditions for all classes and both subjects. It may be noted that in reading, response changes greater than .08 occurred 4 times while in math the response

Figure 5



DATE
DATE
DATE
DATE
PROJECT NO.
4
8
12
16
20
M W F
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
Regular Assignment
Announce Teacher Exchange
Teacher Exchange
Regular Assignments
.89
1.2
1
.72
.099
.082
.17
.087
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
PROJECT NO.
Correct Reading Problems
MOVEMENT
Kunzelmann
Hulten
9 children
14-17
Secondary
TRAINER
ADVISER
MANAGER
PROTEGE
AGE
LABEL

Figure 6

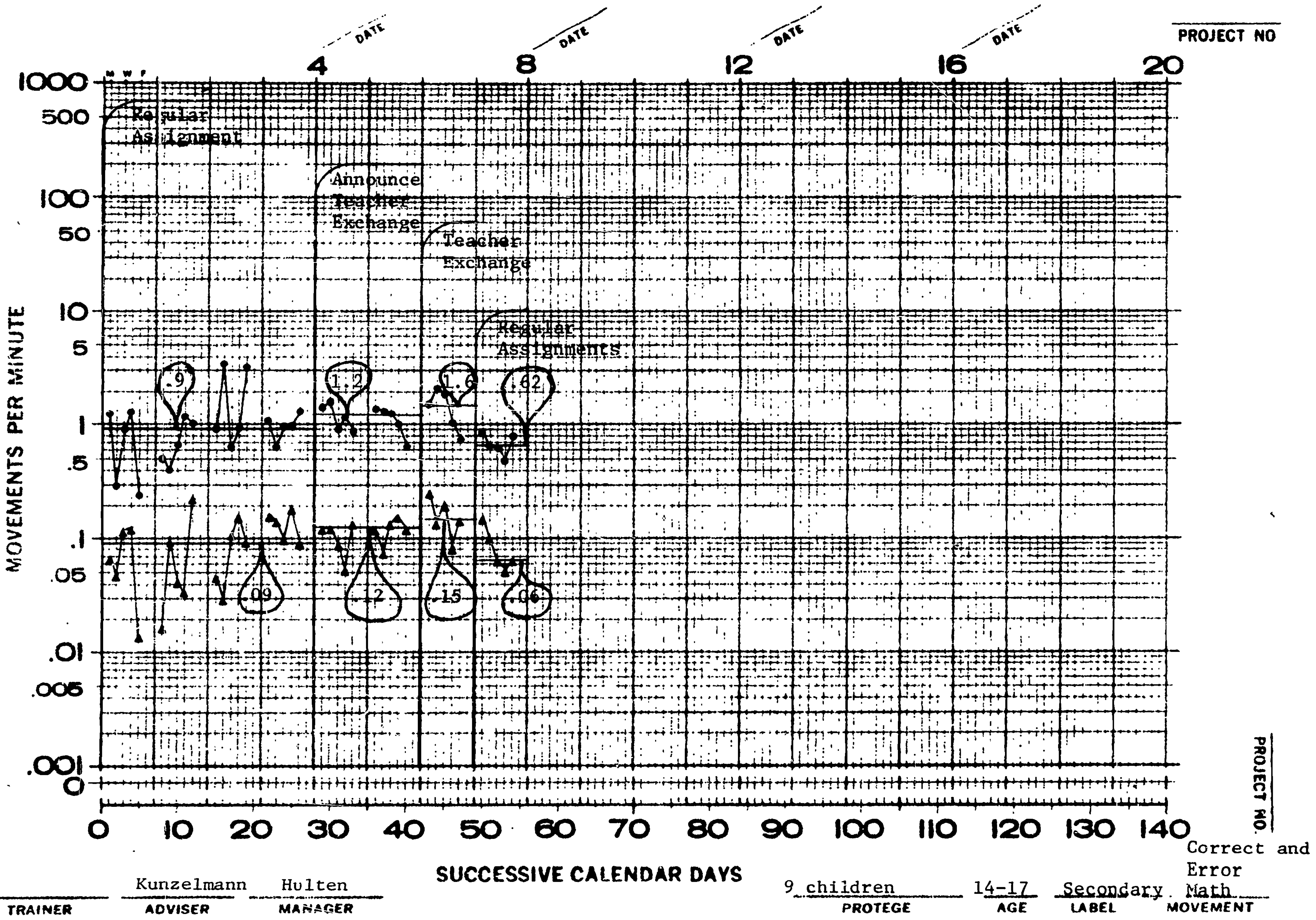

PROJECT NO
DATE
DATE
DATE
DATE
Regular Assignment
Announce Teacher Exchange
Teacher Exchange
Regular Assignments
.9
1.2
1.6
.62
.09
.2
.15
.06
MOVEMENTS PER MINUTE
SUCCESSIVE CALENDAR DAYS
PROJECT NO.
Kunzelmann
Hulten
9 children
14-17
Secondary
Correct and Error Math
TRAINER
ADVISER
MANAGER
PROTEGE
AGE
LABEL
MOVEMENT

Figure 7

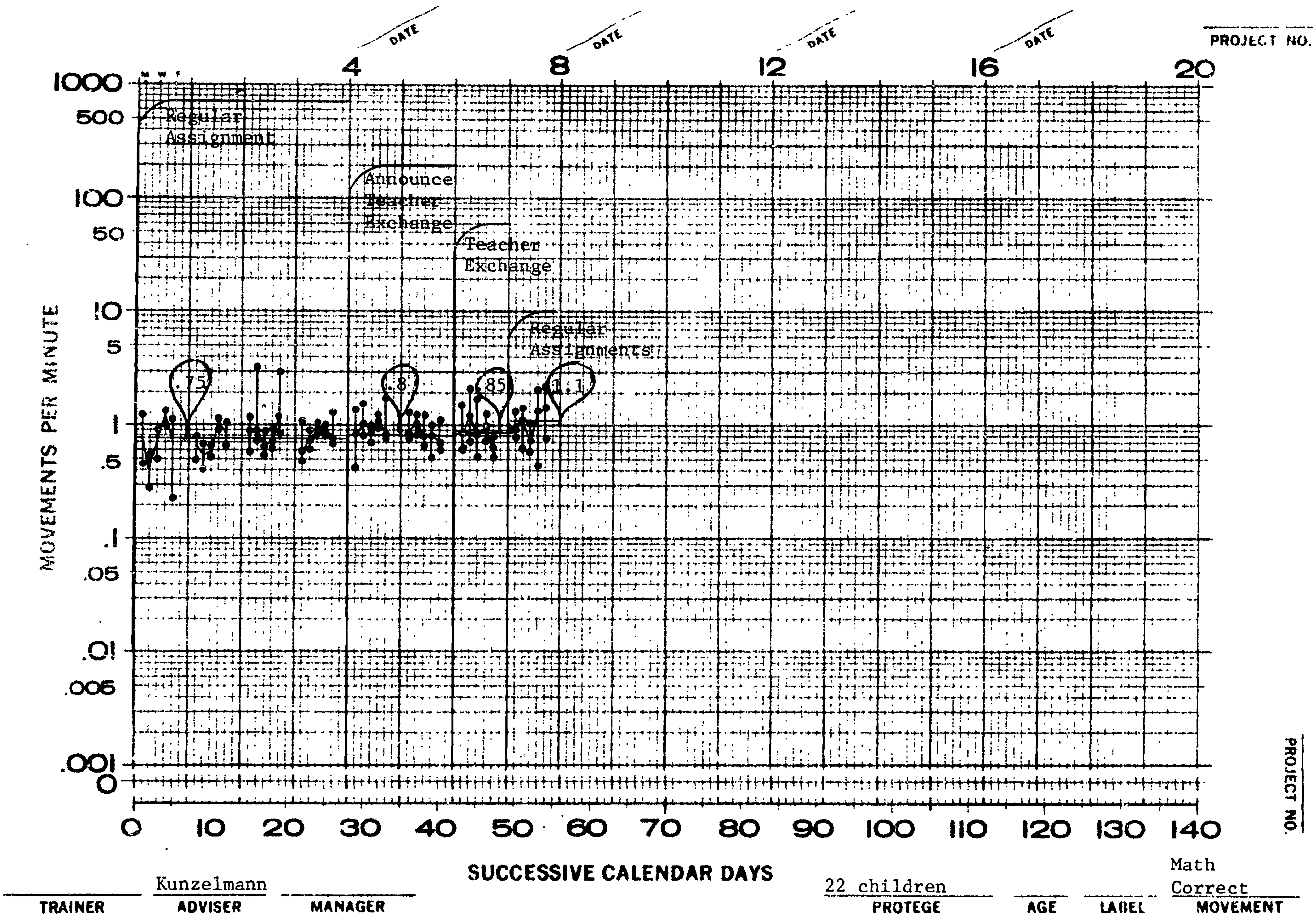
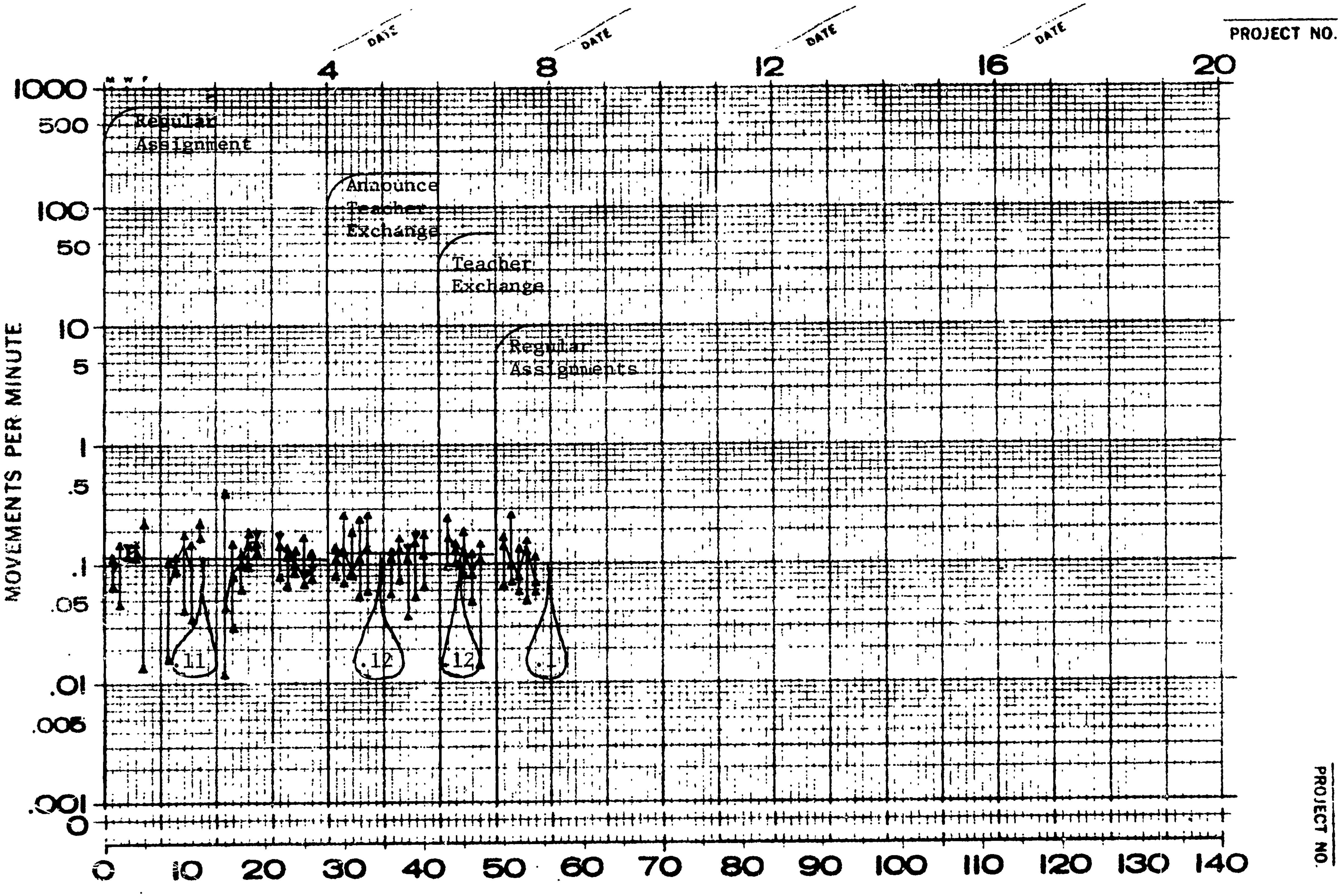

PROJECT NO.
DATE
DATE
DATE
DATE
4
8
12
16
20
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
Regular Assignment
Announce Teacher Exchange
Teacher Exchange
Regular Assignments
.75
.8
.85
1.1
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
TRAINER
Kunzelmann
ADVISER
MANAGER
22 children
PROTEGE
AGE
LABEL
Math Correct
MOVEMENT
PROJECT NO.

Figure 8



Kunzelmann

TRAINER ADVISER MANAGER

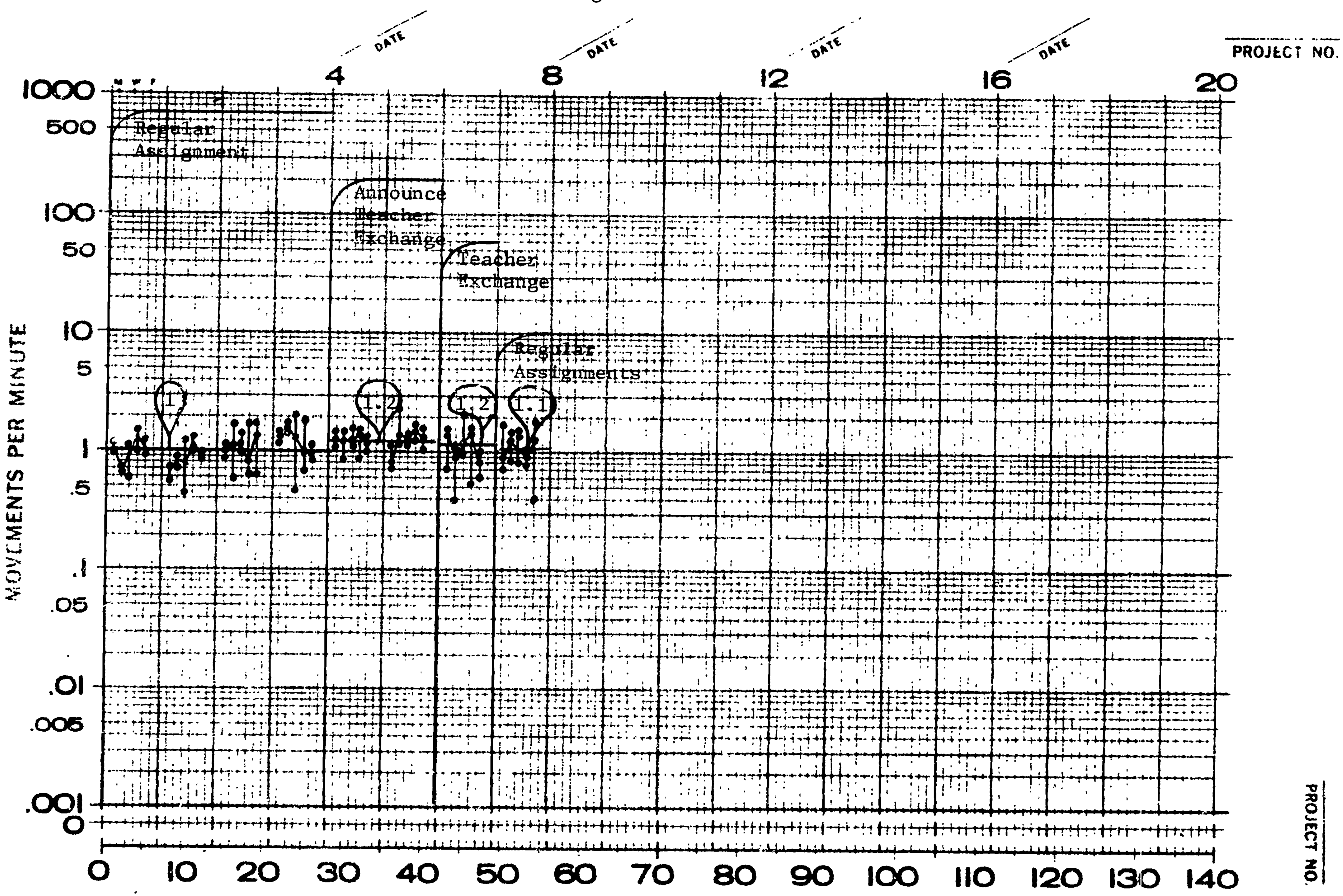
22 children

PROTEGE AGE LABEL

Math Error

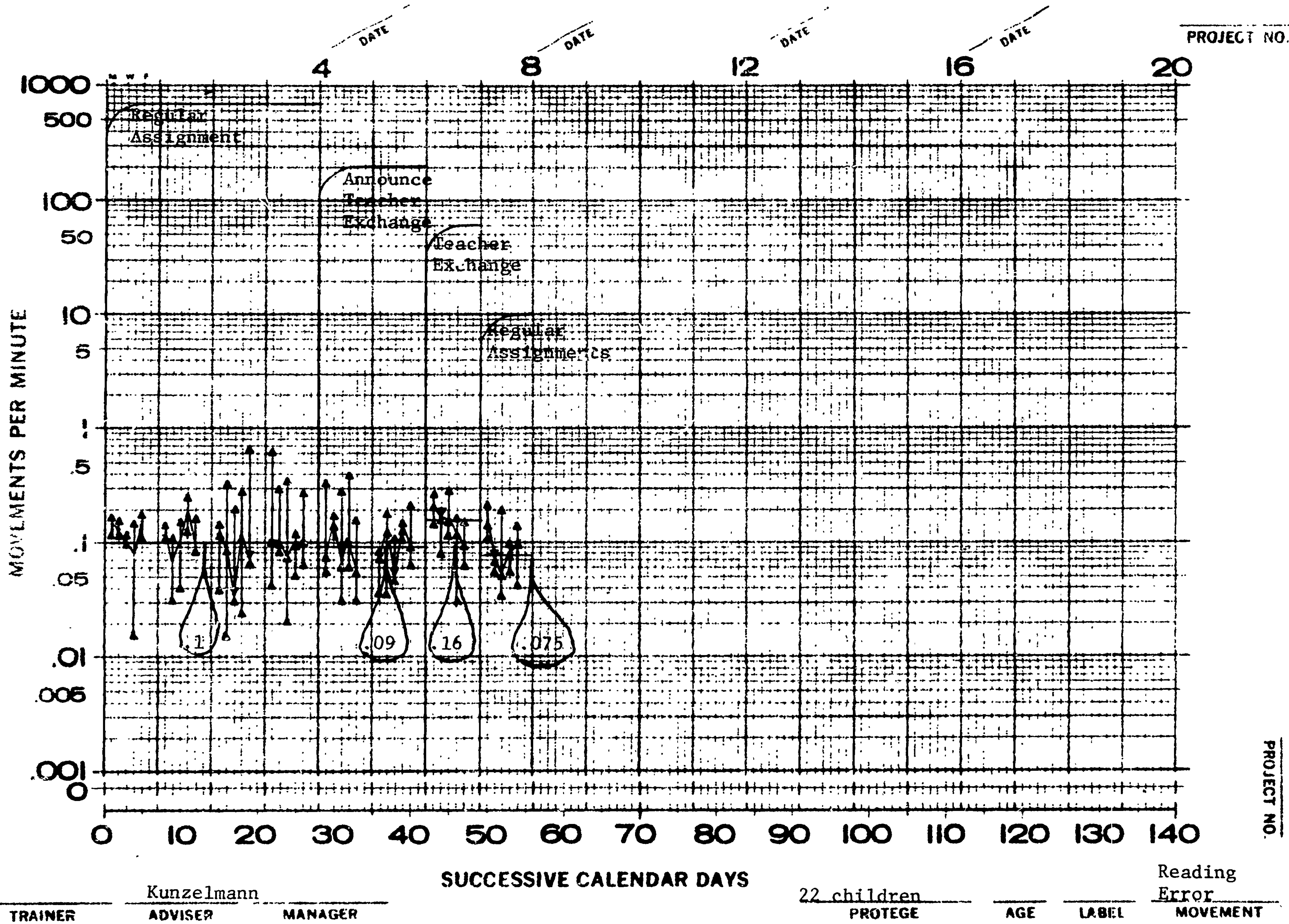
MOVEMENT

Figure 9



TRAINER | Kunzelmann ADVISER | MANAGER | 22 children PROTEGE | AGE | LABEL | Reading Correct MOVEMENT

Figure 10



PROJECT NO.
DATE
4
8
12
16
20
1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001
0
MOVEMENTS PER MINUTE
Regular Assignment
Announce Teacher Exchange
Teacher Exchange
Regular Assignments
.1
.09
.16
.075
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
SUCCESSIVE CALENDAR DAYS
Kunzelmann
TRAINER
ADVISER
MANAGER
22 children
PROTEGE
AGE
LABEL
Reading Error
MOVEMENT

TABLE I

MEDIAN CORRECT RATES PER CLASS
READING AND MATH FOR FOUR (4) CONDITIONS

Class	Sub	Reg Ast	Anou Ex	Ex	Reg Ast
PPI	R	1.0*	1.2	.9	1.2
	M	.7	.8	.68	1.2
INT	R	1.1	1.1	.9	1.2
	M	.82	.8	.8	1.2
SEC	R	.8	1.2	1.0	.72
	M	.9	1.2	1.6	.62

* Rates per Minute

TABLE II

MED. P - VALUES BETWEEN CONDITIONS FOR 3 CLASS
2 SUBJECTS EACH

	Pri-Rd	Pri-M	Int-Rd	Int-M	Sec-Rd	Sec-M	
1 Reg Ast / Ann Ex	.07 Acc.*	.03 Acc.	.3 ---	.3 Acc.	.02 Acc.	.2 Acc.	Direction
1 Reg Ast / Ex	.3 Dec.**	.3 Dec.	.35 Dec.	.35 Dec.	.5 ---	.04 Acc.	Direction
1 Reg Ast / 2 Reg Ast	.1 Acc.	.01 Acc.	.3 Acc.	.3 Acc.	.28 Dec.	.04 Dec.	Direction
Ann Ex / Ex	.1 Dec.	.2 Dec.	.08 Dec.	.4 Dec.	.4 Dec.	.25 Acc.	Direction
Ex / 2 Reg Ast	.24 Acc.	.004 Acc.	.05 Acc.	.36 Acc.	.1 Dec.	.1 Dec.	Direction

* Acc. = Accelerate
** Dec. = Decelerate

changes greater than .05 occurred 5 times. The table reveals that between the regular assignment period and the announced exchange period acceleration occurred in all classes but only in one subject in one class, reading in the intermediate class. Between regular assignment and exchange, a decrease in rate is noted in all but the secondary class where a marked increase in math occurred. Acceleration-deceleration patterns held for all but the secondary program where 4 of 10 possible direction changes did not follow.

Table III is a visual display of the range of data between each condition. This seems important because medians were plotted originally and differences do not reflect highs and lows of pupils per day. An example of more variability would be shown in condition announced exchange and exchange where the increased fluctuation held in all but two subjects where "sameness" was viewed. Data control patterns may be revealed by closer inspection of exact contingencies.

Discussion

The investigation of the recording and data gathering system used for continuous assessment in the Experimental Education Unit classrooms has been started. The data indicate that changes in rate do not usually occur for a class member when fixed contingencies are employed by someone other than the regularly assigned teacher. As Table II indicates, in the second condition, regular assignment and exchange, only in one class, secondary, and in one subject, math, did a change occur that may need replication. The same conclusion also holds for the condition of announced exchange and actual exchange where a decreased in reading at the intermediate level was evident.

The data has opened new questions and more analysis possibilities. For example, pupil rate change per day could produce a finer analysis when based on each individual rate. Secondly, an analysis of the relationship or association of correct and error rates may produce new questions about any of the six contingencies now recorded. A visual picture may be obtained of this association by turning each figure 1-6 sideways for viewing. Note may be made that correct and error rates have a narrower spacing for reading during the exchange condition as discussed in the results section.

TABLE III

VISUAL - VARIABILITY OF MEDIAN RATES

PER CLASS PER SUBJECT PER FIVE CONDITIONS

	Pri-Rd	Pri-M	Int-Rd	Int-M	Sec-Rd	Sec-M
1 Reg Ast / Ann Ex	(-)*	(Same)	(Same)	(+)	(-)	(-)
1 Reg Ast / Ex	(+)**	(Same)	(+)	(+)	(Same)	(Same)
1 Reg Ast / 2 Reg Ast	*** (Same)	(+)	(+)	(+)	(Same)	(-)
Ann Ex / Ex	(+)	(Same)	(+)	(Same)	(+)	(+)
Ex / 2 Reg Ast	(-)	(+)	(-)	(+)	(Same)	(-)

* (-) = Less Variability in Median per day Rates for Class
** (+) = More Variability in Median per day Rates for Class
***(Same)=Appro. Variability in Median per day Rates for Class

Implications

Problems in administration of Special Education classes are generally not identified because of the lack of techniques for assessment. The study shows clearly that more control occurs on pupil performance rates when a change is forthcoming than when the predicted change actually takes place. Although many points were made that exemplified possible problems when teachers were absent from classrooms, the most relevant problems were addressed to lesson planning or clear contingency issues. This study shows that, under conditions of extensive planning for contingency management, changes in teachers do not significantly alter childrens' performance. The findings, no change, have relevance to at least two issues pertinent to administrators of Special Education classes. First, a teacher, trained in precise managerial techniques, can handle many children with group sub-divisions directed by assistants. Materials, desired performance levels, and contingency plans could well be handled by a master teacher. The teacher would have as a major task the analysis of data and directions for environmental changes.

Second, the six independently controllable variables discussed in the procedure section of this paper are variables with enough strength to open the way to exploratory use in special classes. The lack of pupil change when teachers are exchanged should eventually lead to questioning the research on teaching which attempts to correlate teacher personality and teacher effectiveness. Teacher personality was not one of the six variables discussed.

In summary, the tactic of exchanging teachers in Special Education classes may bring about more precise lesson planning and consequently more pupil and teacher control over performance.

CHAPTER XIV EDUCATIONAL ENVIRONMENTS AND ADMINISTRATION

R. J. Kothera

John Galsworthy once said, "If you do not think about the future, you cannot have one."

In any consideration of the future, the present becomes extremely valuable. Speculation concerning what will or might be lacks foundation without the realization of what is at present. The future course of education will not be charted by wishful thinking or total immediate change but by adaption, improving and remodeling our present environment, as well as our behavior, to elicit the goal seeking attitudinal responses necessary to accomplish the objectives of education.

We find ourselves today at another crossroad without an acceptable map by which to plot a course. Unfortunately we must plot a course; we can no longer follow the false leads of the innovator's editorial comment. To change is not only to make things better but also to risk making them worse, to sacrifice sums of money needlessly and, unforgivingly, to perhaps sacrifice children. Within the scope of change lies commitment to research the attributes of contemplated action, to quantify the positive and negative for each alternative, to reach that point of assurance that the definition of educational objectives is adequately approximated and met with firmness, not only of our convictions and allocations, but also with the results of our research and experimentation.

Education is no longer an isolated discipline but one that is crossed and recrossed by several disciplines. Science and technology have great impact in the field of education; future impact appears so great it is incomprehensible. As we now attempt to define the educational environment and its administration, it is of prime importance that we remember that inherent good or evil is not a related aspect of any new development but exists only as an alternative dependent on our convictions and commitments. The possibilities for good are always present; it remains for us to realize them.

In an examination of learning environment, wide discrepancies are found in terms of today's attempts in establishing the interrelationship between the learning process and environmental components. Research increasingly bears out the importance of a controlled environment or an environment that is controllable as a factor, an imposing factor in the learning process. The various environmental components must be considered independently as well as in terms of their interrelationships. These interrelationships are factors not only in the learning process but in the academic program and operate for the benefit or detriment of the overall program.

Physical Environment and Learning

Modern technology presents the architect as well as the educator with a wide choice in design. Most truly modern schools, in those communities willing to take the step forward, are designed utilizing nonrectilinear geometry. Rectilinear, curvilinear, polygonal, and free-form geometry are important in terms of geometry's serving the functions and programs of the educational establishment. They create major areas for instruction or, in another sense, a master module contrived for immediate dissemination of information over a large group, yet inherently capable of bisection to function as focal points of learning for any elements of the large group. Group size then is dependent not upon structure, but upon ability, achievement, interest, behavior, and their relationship to the objectives to be accomplished. The educational program and its objectives control space which exists only for utilization by the determinates of academic progress. All too frequently the program is controlled by nonflexible structures within the space we call a school. If we are to truly meet the needs of children in the academic sense, the spatial component must retain the flexibility necessary to a wide range of abilities and interests inherent in the learner. We have consistently required the learner to conform to the program and the program to conform to the building or to the space devoted to it. We are now entering a period where space is not mere housing but a vital ingredient, a catalyst which facilitates reaction.

If we wish teachers to perform a cooperative effort, then this possibility is enhanced by open space, close association and visual contact with ease but it is hampered by walls, doors, dividers, and self-containment within 32' x 32' x 9' cubes. The majority of schools was designed in terms of these cubes prior to any examination of program or anticipated functions. The effectiveness of team teaching, continuous progress, and ungraded classrooms tend to be diminished because of the inability to break through the physical barriers that isolate classes and faculty throughout the instructional day. Cooperative teaching and planning as well as the adequate placement of pupils by achievement necessitates ready access to appropriate learning situations by direct passage, without traveling those expensive pathways we refer to as corridors. Corridors serve no purpose other than paths. At a cost of $20.00 per square foot, 5,000 square feet of corridor is an expensive item in a building program that really does little except increase the cost of maintenance, heating and lighting. At present architects favor large spatial components adjacent to each other as well as to limited defined laboratory areas for the teaching of subjects with high equipment requirements. Schools of the future should be characterized by space where a multitude of different activities can and do take place. Areas within space will not be defined by subject, level, or grade, but by teachers' names. The names of faculty will relate to special training, talent or special abilities necessary to the learning process as well as to the overall demands of the curricular program. Mr. Beige, for instance, teaches large group physical education courses, 50 to 100 male students per class, but, because he is able to provide problem learners with a father figure and enhance motivation while providing natural social reinforcement, he also spends part of his time with small group study periods besides supervising independent learning and practice sessions in areas of his personal vocational and avocational achievement. The allocations of space for learning are dependent on need and are scheduled on a weekly basis. Students are channeled by weekly schedules into learning suites where contact with teachers is made.

Although space is defined in terms of curriculum and overall program and its utilization determined by individual and various sized group tasks, the learner will also be aided by an imposed physical environment of light, sound, temperature, humidity, color, and aesthetics inter-relating with various organismic factors.

The act of seeing or visualization is a major part of the learning process even though it is possible to learn without this sense. Perception, however, is inherent in learning regardless of the learner's ability to "see" by the use of any sense. Since perception is dependent upon stimulus, the environment must afford the optimum opportunity for stimulus presentation; hence, light or lighting is of major importance. Light, artificial, natural, or in combination, facilitates learning when used properly. Improper usage impedes learning and may result in serious optic deficiences. While a quantity of light is necessary to learning, the quality is of more importance. Glare is controlled by 37 foot lamberts of brightness in the visual field and contrast of light between areas should never exceed a ratio of 10 to 1 and should really be held at approximately 3 to 1. Light should be full enough to easily recognize three-dimensional forms and color controlled to avoid distortion of colored surfaces. What really is necessary in the controlled environment is duplication of natural light which is presently available in the form of fluorescent tubes. You will not receive a sunburn, but it is possible to pick up a few of those necessary vitamins by reading a magazine article in the library under a fluorescent light. While the typical light level is 50-55 foot candles, the intensity of light should vary with the learning situation. Light within a learning suite should be controllable over a wide range, with the circuit box easily revamped by the teacher who wishes to place specific controls, i.e., a dimmer, on a specific section of lights. This would allow controllable intensities in any area of the suite to aid perception. High intense light on a specific demonstration focuses attention and could provide the channel to recognition and perception without repetition. Also of importance is the elimination of visual noise or bright spots introduced by window walls, high brightness light fixtures, or any element that does not contribute to the visual comfort of the learner. Visual noise also competes for student attention by diverting the learner's eyes from planned visual contact, reducing recognition and resolution of audio and visual information.

As we deal with open space schools, the window walls of past classrooms are eliminated to provide more flexibility and less distraction as well as glare. Lighting then becomes more important and its control a prime necessity. Teachers will be able to control the quantity and quality of lighting to effect behavior of the learners within the learning environment.

Color, in windowless space, becomes important as it allows for proper three-dimensional resolution while preventing claustrophobic reactions to color combinations in ratios of high contrast. Colors with a high degree of reflectivity aid in maintaining desired light levels.

Controlling acoustics in the learning environment is not as yet an exact science. We know that groups operating coterminously within a learning suite can do so with little or no acoustical privacy, yet we also know that some attempts to teach in open space have met with disaster. There is the indication that sound itself is not necessarily responsible for the failure of some examples but possibly it is the absence of sound, in addition to teachers who react psychologically to constant visual and auditory contact with administrators and fellow faculty members. The first of these problems may be overcome by introducing white sound at a level of thirty decibels to give an acoustic background to the learning suite while the second may produce data that there are self-contained teachers and open-space teachers with exchange between the two a psychological impossibility. While the ear is a very flexible organ, its human host may lack that one element achieving prominence in today's modern society, flexibility. Acoustics may be further controlled with carpet as well as with tile and plaster. Some schools use carpet on walls as well as hanging drapes. Sound like light must be controllable for use in the learning environment and there is no reason why that control shouldn't be in the hands of the teacher to vary according to the needs of the learning situation.

Temperature and humidity are well researched areas of the physical environment. The desirable temperature for sedentary learning tasks is 72° F. with a relative humidity of 40-60%. Two degrees of temperature rise above the ideal decreases the learning rate approximately 20% and increases above the 74° F. level decrease the learning rate, but at a steadily declining pace. High temperature conditions increase the rate of heart beat and can produce a health hazard that doesn't become apparent until years later. This research should be evidence enough for year-round climate control, not only for students but for employees as well. It has always disturbed me, as a parent as well as an administrator, that the physical environment of the classroom always pleased the teacher while the learner was never consulted, not even through research, which would seem to say to teachers, for a 20% increase in learning, wear a sweater and perhaps a longer dress. Research also shows that many schools have a humidity level of 12-14% during the winter, which is even lower than some desert areas. Extreme aridity is one of the main causes of respiratory complications and aids February and March in remaining at the top of the list in annual death rate.

The structural and functional characteristics of man are called organismic factors that interact with the various elements of the physical environment to determine the environmental levels that will be operable within the learning suite. The age and behavioral pattern as well as sex, diet, and caloric intake and metabolism all are elements which will vary and must be considered in the control of the physical environment. In addition, one other set of factors exists that must als be considered. The reciprocative factors of activity, clothing, degree of necessary exposure, and sociology or group size, movement, and work patterns all affect the alternatives of choice in establishing the physical elements of the learning environment. The physical environment should be mainly in control of the teaching faculty insofar as choices between alternates are concerned; however, administration must ascertain which of the alternates as well as assure their operational level. Yet, individual choices are necessary for teachers; light levels, volume of sound, etc. are constant variables and so should be used not only by the teacher but by the pupil as well where individual control is possible.

Is all this probable? Yes, at the present time and most necessary to create a physical environment that frees the learner from physiological stresses. Dr. Darrell Boyd Harmon, an educational researcher, states that, "An environment free of physiologic stress actually increases mental capabilities of students. A research study performed in El Paso, Texas, (2 1/2 year duration) confirmed a $16.00 savings in the Welfare Budget for every $1.00 spent to improve environmental conditions within school buildings." With the money we would save, let's purchase two items which will improve the learning environment as well as faculty morale.

The first is the invention of Dr. Cristjo Cristofv who discovered a means to end that tired, worn-down, sleepy feeling common to teachers around four o'clock and to students all day. Dr. Cristofv points out that in nature there are plenty of positive ions to keep people pepped up. At sea level this charge equals a field strength of several hundred volts per square yard while in those invigorating mountains the positive ion count may produce a field strength as high as several thousand volts per square yard. Problems appear, however, whenever we enter an enclosed space, such as a building, an automobile, or an airplane. The enclosed space becomes a Faraday cage and its occupants are beset by premature fatigue, drowsiness, and shrinking powers of concentration caused by excess negative ionization of the air. Dr. Christofv's invention, an anti-fatigue device no larger than a shoebox, pours out positive ions which step up the impulse rate of positive electrical charges delivered to the wakefulness center of the brain. The amount of impulses relate to individual alertness. In the living organism, nerve impulses increase the nourishment of cells and delivery of oxygen and heat to tissues. The theory then is: the more impulses, the better you can think. Before we dismiss this as superstition, the Christofv device received U. S. Patent Number 3,311,108 in April, 1967, and is patented in 14 other countries as well. The principle has been used since 1959 in high flying military jets and U-2 reconnaissance planes, by the U. S. in space, as well as in submarines. Work was initiated in 1940 as a Luftwaffe research project to combat the high rate of accidents when its air force switched to enclosed cockpit planes. Cristofv discovered a high negative charge was being created in the cockpit by the plastic covers and seats. This brought swift fatigue and a marked decrease in pilot performance. The device will cost approximately $130.00 and is made by Electrogen

Industries, Inc. at Westbury, New York. With climate controlled, windowless schools, this device could be very important.

The other means of spending that savings to improve the physical environment is to do away with all electric cords, electrical outlets and all the problems associated with machine operation and its metalic umbilical cord that is nourished from some central power source.

Nikola Tesla, a Croatian engineer and former employee of Thomas Alva Edison, developed a successful system of alternating current in 1880 in opposition to Edison's direct current. Thirty years later he was experimenting with a means of sending current out as waves and at one time, in an experiment, lighted miles of light bulbs in a desert area by plugging the bulbs into the ground anywhere he wished light. Unfortunately this genius died before his work was completed and, more unfortunately, his ideas and knowledge died with him. For over forty years others have tried to follow without success, but there is a patent application at present in Washington for passing a current through carpet whereby that current may be utilized anywhere in the building by setting a piece of equipment down to pick up the current for operating the machine. While this isn't about to be marketed in the near future, each of us will see this idea become operational in our life time and probably taken for granted as well. Assuredly, the next step would be the passage of current through the air similar to radio waves.

The Social Environment and Learning

As we enter into a discussion of what the environment will be that directly operates with the pupil, it is possible to define it by the use of the abbreviation ecol. for ecological, which is defined in sociology as "the relationship between the distribution of human groups with reference to material resources and consequent social and cultural patterns." While those outside the educational establishment have often accused us of attempting to establish our own little world, what is being suggested here may indicate that it is necessary that, in part, this be done. The use of the term ecol-system is deliberate. It should emphasize the fact that in education a total establishment must be engendered for, as we view the future, we also view not only scientific and technological change but radical social change. For instance, in the next fifty seven years, we shall probably see:

1. Motivation as the key social problem and definition of public education.
2. Education based on reality, with the past used only as reference.
3. An almost totally diminished pride in state and local government.
4. The use of the term Social-Democracy to describe our country.
5. Fifty per cent of the people will be knowledge workers functioning at sedentary levels.
6. The work week will approximate 22 hours.
7. In order to give better care and more adequate protection, the fetus will be carried outside the body.
8. Everyone will attend school for a minimum of 15 years starting at age three. Teachers will deal with diapers.
9. While university education will be for the professions, local area schools will handle vocational and avocational education on a 16 hour/day basis, 6 days/week.
10. While 15 years of education will be mandatory, at least 50% of the urban and suburban populations will continue education as a recreational activity.
11. Ethnic and sociological differences will diminish to the point where the population might almost be called homogenetic.
12. Total community development will be the plan of the future with schools in television contact with each home as part of the development.
13. Each home in the development will subscribe to the library and information retrieval service which will function to bring any information to the home as well as to fill the role of tax consultant, legal advisor, homework helper, etc. It will be a computer station no larger than a T.V. set.
14. The 15-year educational program will operate as a rehearsal pattern for future real problems and real productivity.
15. Water will be our largest natural resource problem - how to get it clean and keep it clean.

16. Sociological changes will create more problems than technical change.
17. One great sociological problem will be the ever-increasing alienation of urban and suburban societies.
18. All the printed material in the world will be stored in a space the size of a shoebox.
19. Children will be removed from corrosive family environments and placed in dormitory schools where teachers will also function as parents.
20. "Learning by discovery" will only be used by the mentally deficient. Average students will learn by thinking what we are doing.
21. Programmed instruction will automatically individualize 50% of the student day while the other 50% will be spent in independent and large group activities.
22. Group size in education will be established on the basis of interaction studies and group size will approximate 5 to 15 pupils where interaction is required. Where interaction is not necessary, group size may be as large as 300 per session.
23. Found throughout our schools will be "Boob Tubes," automated, programmed learning devices established as carrels for individual work.
24. There will be approximately 1,000 school districts in the 50 United States, with each school in the district a component part of the district research network. The districts will also be united by means of a national network. At the district and national levels educational researchers and research interpretors will function to aid and direct problem solving.
25. All teachers will be under civil service.
26. The learning process will be made as "teacher proof" as possible.
27. Children will be given free choices in education to minimize coercion.
28. Individual study will be the symbol of the "good" student and will be earned through academic achievement, sociological contributions and the spirit of cooperation.

29. Education will carry out the master plan for making segregation and cultural alienation obsolete by utilizing and inculcating in children the processes of:
    a. involvement
    b. commitment
    c. rationality
    d. reality links
30. The means of education will be reality based and problem solutions will be sought only in the area of reality.

This is what we must prepare for.

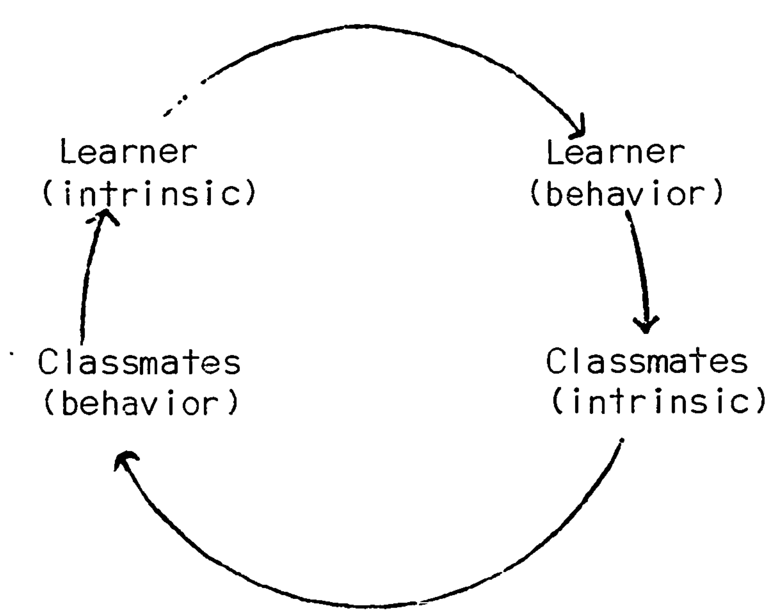
It won't come about with exactness since these are only guesses, educated guesses at best. If the future appears in this light, then education must put its house in order now, to establish an environment with operational aspects which will allow progress without time-consuming and wasteful confrontation. One development that must precede all others is the development of a rational and open mind to social and technological changes in education. We must learn to eyeball, explore, research, experiment, and perforate new ideas and developments from the standpoint of usage and plausibility, rather than from the apex of our prejudice. We no longer have the time for irrational verbiage. Perhaps one element that was left out of the future in the next fifty-seven years is the commonality of the following statement, "Do you think so? Show me your data." In other words, don't tell me about it, prove it with facts. So, in describing the learning environment, let's put it where our data is.

The teacher-learner situation in its traditional but inadequate aspects consists of two major activities, the teacher or information output, and the learner as information input. The learning recipe reads something like this: Take one knowledgeable teacher or one good textbook, add one highly motivated student, and stir continuously disregarding all other ingredients. Result: you may get an Abraham Lincoln or a George Washington Carver. From the thousands of mixtures of this recipe, our country, unfortunately, has received only one Lincoln, only one Carver. Perhaps the others were only half-baked or the ingredients had soured.

Fortunately, an increasing body of research is available dealing with the relationship between learning and social environmental factors. This research challenges our traditional teacher-learner situation and directs our attention toward a managed environment to produce efficient learning dependent on ability for a wide variety of learners.

Learning takes place in an environment involving teacher, parents, community, other students, as well as the intrinsic elements of motivation, values, and the learner's self-concept. We must understand that all these elements, both intrinsic and extrinsic, are recordable, modifiable, and controllable. Perhaps they are not to the degree that would satisfy the purists at present, but we must start now to develop techniques that are adaptable, to find means that are successful, to create demands that will be fulfilled by science and technology, to develop an intelligent, responsible society that will take our work, build on it toward a better end product. Our starting point is the inner processes or the intrinsic elements of the learner.

The learner in the learning situation has intentions to act, to do, to please, to make friends, to be respected, to complete the task, to get good grades. The intentions stimulate behavior from the learner's repertoire which may not effectively accomplish objectives. The intention to study hard may fail due to a lack of study skills or instead of making friends with a display of skill and competence, the learner is repulsed due to too conspicuous a display. The learner's behavior elicits behavior from all extrinsic elements of the social environment which results from the interpreter's interpretation of the learner's behavior in regard to each of the intrinsic elements. They then behave toward the learner who interprets their behavior. This might be referred to as going around in behavioral circles but is best described as behavioral cycles which aid in developing motivation, values, and self-concept.



The behavioral cycle is one of continuous interaction achieving operational balance based on values, motivation, and the self-concept of those involved in addition to skills in the areas of behavior and interpretation. The cycle itself is what we make it, supportive of its membership, or vicious in its attacks upon the intrinsic elements of the learner. The attitudes of all are also shaped by their perception of the attitudes of parents, friends, neighborhood gangs, and teachers which tends to increase the necessity for positive contact between school and community, not only at the school level but at a variety of community levels as well--through scouting, spots, clubs, adult education, and various service organizations. It requires total effort toward objectives.

Schmuck (1962) identified two types of classroom sociometric structures. The first was a "centrally structured" group with a narrow focus of sociometric choices and a "diffusely structured" group with a wide focus. The centrally structured group tends to select the same few members as most liked and least liked while many are never mentioned. The diffusely structured group utilizes a wide variety of "most" and "least" choices, has less chance of developing cliques or subgroups and has few neglected members. These patterns of relationships do make a difference in the learning environment of each child. The learner's relationship with classmates, and his position in peer structure, are related to academic achievement. Zander and Van Egmond (1958) indicate that the relationship is less significant for those with below average intelligence and Schmuck (1962) clarifies this by pointing out that the learner's perception of sociometric status is more important than actual status. Those in centrally structured classes were more accurate in estimating status, due to narrow focus, while those in diffusely structured classes were largely inaccurate and perceived themselves as being highly liked, perhaps due to spread of friendship, wider influence potential, or perceived expertness. The teacher develops a classroom environment that is directed toward a wide variety of opportunities for each learner to be liked, to provide knowledge, or to exert elements of leadership toward objectives.

The learning environment is further enhanced by the teacher's democratic style as described by White and Lippitt (1960):

> The teacher who wants to be "democratic" and also efficient should continually seek to broaden the base of participation in decision making, whenever participation is really functional and not too time consuming; yet he should usually (not always) exert active leadership and he should unhesitatingly, without the slightest feeling of guilt, use his natural authority whenever the situation calls for firm control or for swift, coordinated action.

The relationship between teacher and learner is affected by many factors but Fox, Lippitt, and Schmuck (1964) indicate:

> High satisfaction with the teacher appears to be more powerful than all other social influence factors (social class, parental support, peer status) in affecting full utilization of intelligence.

Generally, the more the learner and the teacher share a friendly attitude. the more they agree in terms of relevant behavior, the more opportunity the teacher has for higher utilization of intelligence on the part of the learner. While Ryan's work, which explored the possibility of matching teachers and pupils proved inconclusive, it does seem to promote more favorable attitudes on the part of teachers and pupils.

Another means of providing a more adequate environment for children and a direct aid to teachers is the use of cross-age helpers (Fox, Lippitt, and Lehman, 1964; Schmuck and Lippitt, 1964). All writers promote in-service training for both older and younger children to overcome problems of attitude and maturity maintained, if not fostered, by the traditional age-graded classrooms of the egg-crate schools. Ronald and Peggy Lippitt, in their journal article of March, 1968, indicate increased emotional and academic support for the younger children, while the older children benefit from reviewing subject matter, receive an opportunity to work through some of their own problems with peers and siblings at a safe emotional distance and make a tremendous effort to fill their own academic gaps when they are responsible for helping someone else understand.

In the analysis of learning, we discover that it is not a social process but really an individual matter. Although the social environment may affect learning, the behavior of the learner takes form and direction through inner processes, a major feature of which is self-concept. Rogers, Maslow, Bledsoe and Garrison, Helper and Coombs all indicate that adequate individuals, self-actuating individuals, have a positive view of themselves. They realize that they are liked, wanted, accepted, and are able. They have worth and importance which resulted from the learner's perception of being liked, being respected, being adequate, and realizing success. Those who perceive themselves as being disliked, rejected, failures of little value have reduced actualization of academic potential. Schmuck, Luszki, and Epperson state this is true regardless of the sources of rejection.

An effective learning environment would provide for enhancement of the learner's self-concept by recording success and failure experiences, by monitoring the kind of feedback children receive from teachers and others and providing a social organization in the school which is more diffuse than central, for as the peer structure increases in diffuseness, the attitudes of learners toward themselves become more positive, not only toward themselves, but toward others and toward their own work.

Yet learning does take place in a social situation and within that environment occur the various stimuli that influence the perception of the learner's self-concept. For the large majority of students, there is evidence that the learner who perceives himself as well liked by his peers, finds his values and attitudes consistent with those of his peers and teachers, feels valued and supported by teachers and parents, and has opportunities to relate positively to children both older and younger, will develop a positive concept of self and will have more potential for the full utilization of his abilities. The learning environment, to maximize opportunities for the learner, will have to be shaped, modeled, and controlled through the use of behavioral modification techniques. These techniques, so successfully used with children having learning disabilities, must of necessity be considered a part of each teacher's techniques in controlling the learning environment for each group. The techniques used in behavioral modification enable the teacher to establish behavioral patterns for students which tend to make them more acceptable to the group since they become more functional in learning and make them more acceptable to themselves since they experience a higher degree of success with classmates and teachers. The work in this area in the Roseland District in Shawnee Mission, Kansas, exemplifies the concept of utilizing behavioral controls to bring about auspicious behavior changes in problem learners, thus eliminating the problem and affording the students more successful opportunities for acceptance, not only on a personal but on a group basis as well. It means that teachers will have to define behavior, find its incidence, level consequences on the behavior, and reward planned changes that occur. This, necessary for individual students, can also be employed for groups, which we see being accomplished daily in the field of special education. Teachers must consider themselves not just information machines but masters of techniques and strategy guiding students into acceptable patterns of behavior to promote greater realization of their potential and greater social and personal acceptance.

Testing achievement, attitudes, interests, social status, etc. furnishes us with valuable grouping information for matching teachers with pupils, pupils with pupils, cross-age helpers with teachers and pupils, and teachers aides with all these. This is easily accomplished through E.D.P., or Educational Data Processing.

The recording and retrieval of data and the control of the learning environment will perhaps be the basic function of teachers once high quality individualized programs are developed. While test data will self-record in each child's assigned memory bank, other data will have to be fed as inputs concerning behavioral controls, acceptable consequences, rejected consequences, and comments concerning program. Attitudes displayed in large and small group activities are most necessary if we are to deal with individual needs.

Instructional methodology must be nongraded and individualized to the fullest extent possible with employment of diffuse groupings dependent on the structure of the building and the applicability of the subject matter. So far nongradedness has had more effect upon administration and organization. Its major interest must be the individual and his differences and its specific purpose, in the words of Frank Brown is:

> . . . coupling the material that the individual is expected to learn with his already acquired knowledge . . .

The curriculum we teach today will be the curriculum of tomorrow but the changes will be found in how we teach what we want the learner to absorb and what we do to speed the process of learning by utilizing behavioral techniques to modify unsuccessful learning behavior.

The teacher's role in the learning environment will be greatly expanded, based on a higher degree of preparation, professionalism and knowledge concerning behavior, learning, and the exact abilities and achievement of each pupil. It is the teacher who will control each learner program and progress, who will plot the size of groups, arranging each to offer the most auspicious consequences to participants. It is the teacher who will decide with the learner, in addition to guiding him along various branches of the curriculum. It is the teacher who establishes the modular program for each pupil and works cooperatively with fellow faculty members preparing programs, evaluating progress and behavior, staffing each pupil each month for necessary changes in program, groups, independent study teachers, teacher tutors or other teachers. Teachers will be employed in accordance with their training and demonstrated ability. Test data and various evaluations concerning personality, avocation, interests, special abilities, and especially, the ability to work with team members both professional and paraprofessional will also aid in employment selection. Teachers must be able to utilize information retrieval systems, perform system analysis in programming, and knowledgeably analyze learning and behavior problems in order to write curricula prescriptions for each student.

The student will be treated first and foremost as an individual and, once his level of achievement is ascertained, will progress through the curriculum at his own best rate regardless of what other students may be doing or what their rate of achievement and performance is. His progress will be based on a standard curriculum with amounts of individual study dependent upon his ability, performance, and motivation. His program may be extremely broad academically or very narrow, and he will determine its scope dependent on approval of his consulting, guiding, or master teacher. He may work with as few as four or as many as fourteen adults, dependent upon him, and, if he has certain abilities or motivations, he may spend 10% of his time as a cross-age helper working and helping younger children overcome curricular problems he already has passed.

It is important not only for students but for teachers to remember that the verb "to learn" is transitive and there must be some "thing" that the student learns. Unless the "thing" seems relevant, the student will have little interest in learning regardless of the excellence of the learning environment. Additionally, there is a growing body of knowledge indicating that different kinds of students require different kinds of teaching strategies. There are also differences in cognitive styles that must be considered since some students require visualization for understanding, while others seem to learn more easily by auditory methods.

The Administrative Environment in Education

As we turn now to the final phase of the environment, it behooves us to re-examine the picture that has been forming of what education could or should be like. The school buildings themselves will be somewhat extreme compared to what is occurring today. They will contain large open spaces which we have been referring to as learning suites and the various elements of the physical environment will be controllable in two ways: first, through the use of cybernetics, such as thermostats; second, through the individual control of teachers as they arrange the physical environment to suit the needs of their stylized academic programs. If anything, the buildings will contain tremendous flexibility in an attempt to answer the needs of the program and its learners. The faculty will operate in a highly cooperative way and it will not be unusual for a faculty member to have responsibility for 100 or more students. The responsibility, however, will not be total in the sense of teaching contact continuously throughout the day, but will be functional in the sense that there will be impingement of the teacher's personality, ability, and guidance throuout a modular week. Faculty will be extremely well trained, chosen with the utmost care and blended together, one with another, to see the most auspicious cooperative result in terms of establishing a social environment to aid the learner in the full utilization of his potential.

Obviously the operations of this school, as well as the operations of the large district, will fall into the hands of administration who unfortunately will have the problem of attempting to displace bureaucratization to the fullest extent possible. This not only endangers functional operations in terms of relationships, but also tends to displace the professional rights and abilities of teachers. It will be impossible, however, to operate large school districts without some form of bureaucracy and without some form of representative faculty committee, functioning to overthrow bureaucratic rule and bureaucratic inefficiency. The larger we become, the more complicated the administrative functions and the greater necessity for more administrative positions. In the future advanced age, cybernetics could quite easily become the one item which will make extremely large districts workable with the immediacy we now find in the rather small district. Administration will function and serve as gatherers and collectors of data, purveyors of policy, and decision makers. Let us examine three functions of administration: the hiring of personnel, the definition of program, and the allocation of resources as three administrative functions found in the schools of the future.

For approximately the last ten years, school administrators, to a small extent, and the university department of educational administration, to a large extent, have been hunting for a theory of administration. We have had a number of theories but perhaps the one that offers the greatest return is Getzels' Model or theorem which he developed at the University of Chicago. It has perhaps been the most lasting of administrative theories. Getzels' Model is based on the social system and he indicates with a model that in the social system there are two dimensions, the nomothetic and idiographic.

Nomothetic Dimension

Social System → Institution → Role → Expectation → Observed Behavior

Social System → Individual → Personality → Need-Disposition → Observed Behavior

Idiographic Dimension

The nomothetic dimension basically describes the institution, the role the employee is expected to play, and the expectation of the employee in terms of organization, which all tend to determine the observed behavior of the employee within the organization.

The idiographic dimension is basically the personal dimension of the individual, his personality, and the need-disposition area, or his personal, individual needs. These make up the individual side of organizational man or the personnel side, and also determine the behaviors which we will observe.

The administrative processes were hypothesized by Getzels to be dependent on the nature of the overlap of the perception of expectations concerning the subordinate and the superordinate. In other words, the superior has expectations of the role the subordinate is to play and the subordinate also has expectations concerning the role he is to play. As long as there is agreement in role expectations there is little conflict. Conflict enters the picture when disagreement occurs between the subordinate and superordinate as to what the role really is. In addition to this area of conflict, there is also role personality conflict where the intrinsic elements of the individual do not allow him to carry out the extrinsic elements of his role. The model has worked well in research and has been particularly useful in studies of dissatisfaction among personnel. In the employment of personnel it is extremely necessary that:

1. They know the roles they are expected to play.
2. Their personalities will allow them to play that role.
3. The expectations they have as to their rewards are realistic in terms of the rewards the organization is able to give.
4. The needs-disposition of the individuals must of necessity be satisfied in terms of rewards and role expectations within the organization.

To put it in the terms of the behaviorists, for organization to elicit high frequency behavior, it must also establish consequences favorable to the organism. Positive reinforcement on either a fixed or intermittent schedule is necessary to maintain the behavior. Personally, we may not like what I have just said, but frankly, we are no different in many respects from any other organism. This doesn't mean to imply that teachers will perform an act conducive to education and a light will turn on and a handful of peanuts will appear in their mailboxes, but I do mean to say that administration is going to have to realize that teachers are human with human needs and human desires and that as we talk about satisfying the needs of our students, we must also talk about satisfying the needs of the professional faculty as well as all employees.

There are more teachers and administrators who are interested in rewards other than money and these rewards tend to deal more with personality and needs-disposition than with finances. This doesn't mean that money is not important. It is extremely important, but only important to a point, and beyond that point the satisfaction that we derive from our work becomes more important than additional dollars. There is a school district which does not offer the highest salaries in the area but tends to bring in highly creative, aggressive teachers who want to perform in education. They come to work for this district at less money because they are offered freedom to work, freedom to try out ideas, freedom and encouragement and financial backing to motivate, to create, to seek greatness in their own right. Administration must answer these needs. Many times the reinforcement is nothing more than, "Gee, Clara, you really do a good job." People like to hear how they are doing. They like that offhand remark that tends to be an evaluation, especially if it is positive. At the same time, we must also look at the area of negative reinforcement as a possible means of getting a resignation without having to fight state and national organizations and the teachers' associations in order to accomplish the removal of a teacher who is not satisfactory or not even attempting to meet the needs of the children. We must also look at extinction in the same way. Getzels' Model at present can be used for the employment of personnel, for maintenance and promotion of personnel, and for the discharge of personnel.

As administration and professional faculty work together to define and redefine the program and its applicability to methodology, we shall find the problem of experimental time to be one of the greatest hindrances to change in the curriculum. By experimental time, I mean that year or two in which one class is established as an experimental one and another class acts as the control to test an element of the curriculum, gather data, and interpret it so that a decision can be made. One group are the "haves" and the other group are the "have nots." At the end of the experiment we must realize that whichever way it goes, half the children did not do as well because of us. Basically, half the children wasted time that could be spent to better advantage. This is why the development of systems analysis holds so much potential for us in education, for what is done in 2 years can now be done in 1/10th, 1/100th, or 1/1000th of the time, and by the time the hardware is completely developed for us to utilize, 1/1,000,000th of the time, and no one will be hurt. No one will have been misused because administration and professional faculty attempted to find out the potential of alternatives in curricular decisions.

Now the word "systems" is quite popular in the literature on electronic data processing. To some people it is almost a reverent word and it is used quite broadly as systems review, systems analysis, systems approach. Basically, the systems concept is used today quite widely in business management as well as in engineering and now, finally, in education, the step-child. Systems analysis is a way of dealing with the future and coping with the complexities of modern education. Its feet are firmly planted in the examination of measures that must be performed with a comparison of the alternatives available for action. This approach relies heavily on quantitative analysis techniques and is involved with the alternatives of relating ends to means so that administration can make decisions with a clearer idea of the alternate choices and potential results. A system is any complex unit formed of diverse parts that is subject to a plan or serves a common purpose. Synonyms for the word system are scheme, network, and organism. In educational data processing, the word system implies a systematic look at related operations in an attempt to find a pattern that unifies these operations as well as both logical and efficient operating procedures. As we re-examine systems again, it is an area of resources that has been designed and allocated to achieve an objective according to a definite plan of action.

Administration has hereditarily been the decision maker and only in a few leading school districts has faculty involvement been exemplary. With the growth of extremely large school districts and the growth of extremely large faculties, decision making as become so complex and will become even more complex that information retrieval will be almost an impossibility unless it is computerized. This gives rise to computer analysis of educational programs and problems and the use of simulation in education upon which to base the choice of alternatives in decision making.

Egbert and Cogswell (1964), in their work for the System Development Corporation, have applied system analysis to educational problems of a continuous progress plan high school and, although no such high school exists, they have worked out what they call an auspicious sequence. Each element of the academic program as well as the social environment is given a symbol and plotted out in charts to show flow patterns of operation. Perhaps a brief quote from their report will best give us an idea of the use of systems analysis in programming students. Each symbol represents a computer operation.

> After a student's scores have been obtained (2AØ1), the verbal scores are examined (2AØ3) if an expectancy level is being established in a verbal area (2QØ3), and the quantitative scores are examined (2AØ4) if an expectancy level is being set for a quantitative course (2QØ4).
>
> If the scores are skewed badly toward the positive end of the distribution (2QØ6), a score half-way between the mean and the lowest score is recorded (2AØ6). If the scores are distributed approximately normally (2QØ7), the mean score is recorded (2AØ7). If the scores are skewed badly toward the negative end of the distribution (2QØ8), a score half-way between the mean and the highest score is recorded (2AØ8).
>
> If the recorded score if 36.2 or below (2Q1Ø), the student is assigned to II group (2A1Ø). If the recorded score is from 36.3 to 44.7 inclusive (2Q11), the student is assigned to K group (2A11). If the recorded score is from 44.8 to 55.2 inclusive (2Q12), the student is assigned to E group (2A12). If the score is from 55.3 to 64.7 inclusive (2Q13), the student is assigned to T group (2A13). If the score is higher than 64.8 (2Q14), the student is assigned to O group (2A14).

Here we see a map of the assignment of students through a computer based upon achievement levels. A system of 1,000 students could be programmed in less time than the average coffee break. The technique and methodology of this concept are totally foreign to us now, but we must endeavor to bring up our knowledge backgrounds to cope with these possibilities for they lie in the very near future. As education moves toward a more sophisticated way of dealing with the learner, it is becoming more complex. The complexity is necessary. The learner is a complex organism and if we truly mean what we say when we declare that we wish to deal with the individual needs of students, then basically we deal with the most complex situation of all.

Payroll, management, inventory control, and accounting all illustrate administrative application in systems analysis or educational data processing. The computer can be programmed for each of these tasks with ease; however, to use a computer for such pedestrian performances really limits the attainment of its full potential. If administration appropriately configures E.D.P., it can handle fantastically large volumes of data with no difficulty in relatively short time spans. The sophistication of individualized instruction in a wide variety of fields for a wide variety of learners is possible with computers. Computer assisted instruction, hardware-wise, is extremely simple; program-wise extremely difficult. The student utilizes a typewriter like a keyboard for answering or continuing communications with the computer. Computer output is printed on paper or displayed on a T.V. tube. An audio device may supplement the screen since audio responses appear to be particularly crucial to primary and intermediate age children. Presently in the United States, there is no validated, useable computer assisted learning program that can compete with live teachers in an economical way, but the development of computer assisted instruction is proceeding at a fast pace and eventually, within the lives of each of us, will not only compete with live teachers but will surpass their ability to present information to a large majority of students. The basic problem here is the design of the details for presenting the content in the hardware involved. This immediately raises an issue. Can learning become a process controlled completely on the basis of pre-established criteria operating in an information decision system without human intervention? There are those who see the teacher replaced by computers and imply that learning can be a completely automated process. Those who take a less extreme position see the computer as another tool of the professional practitioners in education. We must always remember one important item. Computers will never be a substitute for professional judgement. They may modify approaches to learning. They may free the teacher from tiresome tasks. They may enable administrators to reach decisions with speed and accuracy, but nowhere is seen the reduction of need for personnel. If anything, new techniques and methodology should free the teacher from acting as a fact presenter and bring into teaching the more professional expectations of the teacher functioning as a diagnostician of learning problems. Systems analysis, computers, and educational data processing should be looked upon for what it is, a technological advancement that expands the professional capabilities of the teachers and administrators.

Administrators will use computers as information systems. There is no end to the type of information that can be stored and retrieved with this equipment. You can take any element of a program, analyze the way it works, what it contains, the benefits it offers, and receive data comparing this program with any other as to the overall productivity of any given aspect. You can seek solutions in the choice of textbooks in terms of productivity on the part of students as well as any aspect of material used just so long as objectives are defined as well as the means by which you evaluate the programs. It is possible to find out instantly the exact location of any given student or teacher at any given time, either their physical presence, their academic placement, their health, or their displayed behavior in a wide variety of circumstances. E.D.P. can make the job of programming students easy and complete. Programming students into new educational programs will be one of our most difficult functions and if the students are ever free from age-gradedness, an impossible one to perform, without the use of educational data processing.

This brings to mind a problem and that is the freedom with which a computer, once programmed, can put out sensitive data. Sensitive data includes disciplinary actions, home and family data that can become quite intimate, financial data, and items which now constitute privileged information. While this has always been a concern in public education, it is greatly magnified through the use of E.D.P. Since confidential data on individuals or groups will be stored in this equipment, the old stalling tactic which all have used of not having enough time or "it's too costly to compile the information," no longer applies.

Another item of importance to administrators is the increased demand upon them for constant consideration of moral and ethical consequences of policies they choose and implement. The demand will increase as administrators use computers more and more to aid in decision making. The administrator will become an almost professional student, not only in techniques of rationalized decision making, but especially of the humanities. Even more important is the fact that computers aid in decision making only to the extent they are able to put out literally what has been put in. They are rational only in their own terms and rationality as well as strategy and tactics must still remain the forte of the administrator. For example, the administrator may find himself, because of the application of the computer, with a decision by which he has just saved the school district 22 1/2% of its budget, but in the process lost 386 teachers who resigned in retaliation of his irrational judgment. Work will be much easier;

burdens will be much greater. Administration in the future as well as today must keep apprised of school needs, of the media for communicating the needs of schools, and the ways in which schools may initiate action. This job will remain with administrators for they still are and will be the only practical force available. It is a most important task of administrators. For, unless the public is willing to provide funds, all ambitions for children's learning, all competencies of the professional faculty, all of the extremely noble philosophies are left stranded. Some administrators will continue in the position of making expectations possible. They will remain facilitators of new knowledge and methodology, acting as levers, actuators, lubricators, the link between community desire and faculty performance. Administration is the essential mechanism which tends to determine activity. While the decisions will be made that indicate the pathways of the future of education by a cooperative effort on the part of teachers and administrators, the responsibility for success or failure still will remain that of administrators. While teachers will be working in cooperation with each other, the same is true of administrators as they weld themselves into a team of specialists to perform their functions with more adequacy than is presently visible. It will be important that each school system has enough administrators to look after the widely divergent expectations and varied performances that go into this recipe which is called education and out of which is obtained a product that, in having its needs well served educationally, will continue into adulthood to serve the needs of our nation at large.

References

______. E.D.P. and the school administrator. AASA Committee on Electronic Data Processing, American Association of School Administrators. Washington, D.C., 1967.

______. SER 2, environmental evaluations. Architectural Research Laboratory. Ann Arbor: University of Michigan, 1965.

______. SER 3, environmental analysis. Architectural Research Laboratory. Ann Arbor: University of Michigan, 1965.

Bledsoe, J. C. and Garrison, K. C. The self concepts of elementary school children in relation to their academic achievement, intelligence, interests, and manifest anxiety. Final Report, Cooperative Research Project No. 1008, U. S. Office of Education, 1962.

Brown, B. F. The appropriate placement school: A sophisticated nongraded curriculum. West Nyack, N. Y.: Parker Publishing, 1965.

______. Black box that turns you on. <u>Business Week</u>, May, 1967.

Coombs, A. W. A perceptual view of the adequate personality. <u>Yearbook 1962</u>, Chapter 5. Washington: Association for Supervision and Curriculum Development, 1962.

Egbert, R. L. and Cogswell, J. F. <u>System design for a continuous progress school</u>. Santa Monica: System Development Corporation, 196[illegible].

Fensch, E. A. and Wilson, R. E. <u>The superintendency team</u>. Columbus, Ohio: Charles E. Merrill, 1964.

Fox, R. S., Lippitt, R. O. and Schmuck, R. A. Pupil-teacher adjustment and mutual adaptation in creating classroom learning environments. Final Report, U. S. Office of Education Cooperative Research Project No. 1167. Ann Arbor: Institute for Social Research, University of Michigan, 1964.

Griffiths, D. E. <u>Administrative theory</u>. New York: Appleton-Century-Crofts, 1959.

Lewin, K., Lippitt, R. O., and White, R. K. Patterns of aggressive behavior in experimentally created social climates. <u>Journal of Social Psychology</u>, 1939, <u>10</u>, 271-279.

Lippitt, R. O. <u>An experimental study of three social climates</u>. Unpublished doctoral dissertation, University of Iowa, 1940.

Lippitt, R. [illegible] and Lippitt, P. Cross-age helpers. <u>Journal of the National Education Association</u>, March, 1968.

Ohm, R. E. Organizational goals--A systems approach. Paper presented at Conference of Professors of Educational Administration. Indiana University, August, 1966.

Schmuck, R. A. Social-emotional characteristics of classroom peer groups. Unpublished doctoral dissertation, University of Michigan, 1962.

Schmuck, R. A., Lusyki, M. B. and Epperson, D. C. Interpersonal relations and mental health in the classroom. Mental Hygiene, 1963, 47 (2), 289-299.

Shaver, J. A. The learning environment. Journal of School Health, 1967, 37, (5).

White, R. K. and Lippitt, R. O. Autocracy and democracy. New York: Harper, 1960.

Zander, A. and Van Egmond, E. E. The relationship of intelligence and social power to the interpersonal behavior of children. Journal of Educational Psychology, 1958, 49 (5), 257-68.

CHAPTER XV THE NEW IMPACT: TECHNOLOGY IN EDUCATION

T. F. Robertson

In a recent article on computer-assisted instruction, Dr. Patrick Suppes of Stanford University described a typical day in an English grammar school in the 14th century as going something like this:

The teacher would read selected portions of a text from a manuscript, explaining parts of that text as he went along, usually repeating each sentence twice. The students were expected to memorize what the teacher said. The following day, the recitation would begin on what had been covered the previous day. There were no printed books; even manuscripts for use by the teachers were rare. The process of education was almost entirely oral, with a very limited use of visual materials of any sort.

Comparing that situation with what exists today, we are all struck with the extent to which printed material dominates a large part of personal and mass communications; and in particular, the importance of printed material in regular school instruction.

Although the place of books in our society is now widespread, it should be remembered that the printing of books on a large scale did not really get under way until the 18th century.

Now we are in the midst of an information explosion. I would like to cite some statistics. In 1765, there was one scientific journal in existence in the world. By 1800, there were 100; 1,000 by 1850; and 10,000 by 1900. By 1950, 30 separate journals concerned themselves solely with abstracting synopses from the tens of thousands of scientific and technical journals in existence.

It is estimated that man's total body of knowledge doubled between 1775 and 1900, a period of 125 years. Between 1900 and 1950, the great body of knowledge doubled again. Between 1950 and 1958, a short span of 8 years, the huge body of man's knowledge doubled again and is now estimated to be doubling every 5 - 8 years.

An indication of concern with this tremendous increase in information is the recent Office of Education appropriation of $3.4 million in a Library and Information Sciences Research Program to "devise better ways of coping with the information explosion." Under Title II of the Higher Education Act of 1965, grants in amounts from $1,700 to over $400,000 were given for 38 projects by educational institutions, libraries and research organizations.

Because of the rapid rate of technological change that now exists, students must expect to train or retrain for at least three different careers in their lifetimes.

If we were just to continue as we are today, in order to educate the school population, by 1975 we would need to have one college graduate in four become a teacher. Because of the information explosion, rising school populations, and rising demands for education and training, traditional approaches and procedures regarding when, how and what is taught have to be assessed. The roles, responsibilities, and concepts for the teacher, the library, and text materials must, of necessity, change.

There is every reason to believe that a revolution in the educational system (what John Goodlad calls an accelerating evolution) comparable to that of the introduction of textbooks has already begun; and that is the new impact: technology in education.

Though it may serve to change some educator beliefs, educational technology can also help meet the tremendous pressure on the school system. We have known for many years that children enter school with remarkably different capabilities and that they will work at different rates and at different levels of accuracy and understanding. However, for obvious economic reasons, we are not able to offer a curriculum program to each child according to his needs. We simply cannot afford that many teachers - even if they were available.

A major promise of new education technology is that it leads to individualized instruction. The hope is that the technology can make the teaching of each student adaptive to his own needs and capabilities;

AND...

Of all the instruments of the new educational technology, the system which seems to hold the most hope for individualized instruction is the computer.

How the computer can be used in the education process is a subject of much discussion and an area in which a great deal of activity is taking place. I will mention briefly some of that activity as I relate ways in which the computer can and will be used in education.

The obvious use, and the one with which we are most familiar, of course, is in its use for administrative as opposed to instructional purposes. Even in the administrative area, education has not been able to make full use of the computer. We are aware of the money problems and the staff problems to do data processing in educational institutions. The educational community needs people who can bridge the chasm between data processing and education. This is a desperate need. This could be a whole subject for discussion in itself; however, today I wish to focus attention on the use of the computer for instruction.

Since 1955, there have been three major stages of application of computer technology to education:

1. University computing centers.
2. Electronic data processing in accounting and record keeping.
3. Now barely under way: on-line, computer-based instruction.

This latter application spells a revolution in American education, primarily because of the development of simultaneous use of computers using remote terminals. Herein lies the most promise for individual instruction.

Generally, computers could be used in the instructional process as follows:

Computation. This was one of the earliest educational uses of computers and has great promise. A college student in his freshman year could take an introductory programming course to make routine use of the computer in many courses throughout his undergraduate career. Students at the University of Michigan, Carnegie Institute of Technology, Washington University, University of New Mexico, Harvard, and Dartmouth, among others, are making computational use of the computer for undergraduate courses.

Information storage and retrieval. Information retrieval systems will provide up-to-date information in any area of the arts and sciences through techniques of rapid retrieval and dissemination of data. Such systems will provide vital information for guidance and counseling purposes. This will be major use of the computer in education.

Computer-assisted instruction. No term which is computer-related has had so many varied names as this one. We hear, in addition to "computer-assisted instruction," "computer-based instruction," "computer-administered instruction," "computer-aided instruction," "computer-assisted learning," "computer-based learning," and others.

What is computer-assisted instruction? It is an educational innovation in which a student is guided by a computer program through an organized but flexible course of individualized instruction.

By assuming a portion of the total instructional process, computer-assisted instruction provides the teacher with more time to give other classroom tasks, such as guiding the student through the inquiry-and-discovery process and providing additional time for work with individual students.

The content or subject matter written for the computer is similar to a series of textbooks broken into small bits of information to be absorbed by the student. Unlike the textbook, however, computer-assisted instruction can provide lesson material on many levels of difficulty for each grade level.

There is a great deal of agreement as to the various areas of CAI which I wish to explain. The first area, drill and practice, is designed to supplement the regular teaching process. Students are presented with lessons on the concept which the class is currently studying and on ability level which correlates with his past performance.

After the teacher has introduced new concepts and ideas in the standard fashion, the computer program provides regular review and practice of basic concepts and skills. In elementary school mathematics, for example, each student would receive 15 or 20 exercises a day, working by himself at a student terminal. These would be automatically presented, evaluated, and scored by the computer program without any effort by the classroom teacher.

These exercises are presented to the student on an individualized basis, and the level of difficulty for each set of questions is based on the student's performance on the previous lessons. If the student does better than "average", he gets more difficult questions to challenge him; if he scores low, he is given easier questions that he can handle.

A question that is frequently asked is, "How will a computer change the teacher's role?" Drill-and-practice systems will modify the teacher's role only slightly. What they will do is relieve teachers of some of the burden of preparing, correcting, and grading large numbers of drill exercises. Instead, daily individualized drill-and-practice exercises in basic concepts and skills will be given to each student, graded, and recorded in the individual student's file; and the teacher will receive detailed reports of individual and group progress.

The second type of computer-assisted instruction, tutorial systems, takes over the main responsibility for helping the student understand the concept and develop skill in using it (more of a man-machine approach). Basic concepts in various subjects can be introduced by the computer program in such systems. The aim is to approximate the interaction a patient tutor would have with an individual student.

In tutorial CAI, the computer program will provide the basic idea in the concept being presented. This program would be written so that if the student did not understand the basic concepts on the first presentation, he would receive a second, and possibly even a third, exposure to them. These latter presentations could be of a simpler or more fundamental nature.

The teacher will be more significantly affected by tutorial systems. The new role of the teacher would be to work individually with the students on whatever problems and questions they may have in assessing and handling the new concepts. Tutorial systems allow teachers greater opportunity for personal interaction with students.

The third and much more sophisticated level of CAI is variously called dialogue, gaming, or simulation. The student could be involved, for instance, in the simulation of actual situations involving problem-solving experiences. As an example, in medicine simulated medical diagnosis can be performed on the computer terminal. Data concerning symptoms, such as temperature, pupil dilation, reflexes, respiration, blood pressure, etc., would be typed. The computer program would then feed back the results of such inputs. This same approach could be used in dealing with chemistry experiments and trouble-shooting problems involving electronic equipment. Other uses in this mode:

1. Supplying periodic economic or population forecasts;
2. Helping to balance the budget;
3. Giving guidance in the planning of new educational facilities.

The goal of the dialogue approach is to provide the richest possible student-system (man-machine) interaction where the student is free to construct unrestricted natural language responses, ask questions, and in general exercise almost complete control over sequence of learning events.

In actual operation, elementary schools in Palo Alto, Cupertino, Menlo Park, and East Palo Alto, California, have had on-going CAI for up to three years.

In the spring of 1968, students in the nation's largest public school system commenced a CAI program in mathematics in grades 1 through 6 under a Title III ESEA grant. This system, when fully installed, will have as its central computer an RCA Spectra 70/45, which will drive 192 student terminals operating on-line at the same time in 15 schools in the New York City school system, and handle over 10,000 students daily in the public and parochial schools, and adult education.

In addition, CAI pilot projects are being provided in four Philadelphia high schools, for sixth graders in Yorktown, and several elementary schools in San Antonio, Texas.

Two projects of a somewhat unusual nature in computer-assisted instruction are taking place at Stanford University. Dr. Wolfgang Kuhn of Stanford's Music Department is using a computer program to teach pitch to singers. Professor John Van Campen has devised a learning system to teach elementary Russian, relying entirely on computerized instruction to the complete exclusion of the classroom teacher.

Another frequently asked question is, "How about cost?" To be applied on any vast scale, CAI must be economically competitive with other systems performing similar functions. Some increase in equipment cost can be justified by improving teaching efficiency and subsequent reduction in training time and operational waste. It seems clear that computer-based systems must approach the over-all cost of more conventional equipment if they are to be used on a vast scale. The development of miniaturized large-capacity systems is a trend in this development.

Another avenue of cost reduction is through the development of special-purpose computers and associated equipment for special educational application. Such computers could be of highly simple design, with storage capacity, operating speed, and flexibility geared to the application. Such a device could be subsequently tied into a large-scale, multiprocessing miniaturized system and be used as a satellite, with access to the power of the large system. There has been a growing discussion of the "Computer Utility" since it was first discussed in 1961 by Professor John McCarthy of MIT, now of Stanford University.

Basically, a computer utility would be a mechanism for permitting a number of remotely located customers to utilize the facilities of a large central computer complex, with much the same ease and flexibility as if the computer were located on their own premises. Such a utility would be analogous to that of an electrical distribution system which supplies electrical power where and when required to facilitate the performance of physical work. In this case, the computer would supply logical power to aid the individual in solving his processing and computational problems.

In a recent issue of Scientific American magazine, Professor McCarthy stated that household computer terminals connected to public utility computers will some day be as commonplace as home telephones. To satisfy their special needs, the terminal equipment used by engineers, scientists, the educational community, and the public would vary in complexity based on requirements. This could range from a very complex terminal for use in research, to the terminal in the home which Douglas Parkhill of MITRE calls the "Fireside Computer." A vital part of such a system would be the communications links connecting the system. Consequently, in order for such a utility to be possible, the development of wide band width, low-cost data transmission equipment, which would permit, say, direct television transmission for the cost of today's telephone service, would have an important effect on the realization of such a utility.

Obviously another important requirement in such a system would be a low-cost terminal, and the figure that seems to be mentioned most at this time is a device that would cost less than $1,000. A minimum facility at such a console would include: a flexible, bright, viewing screen, which would provide full graphic, alpha-numeric and pictorial output in any desired format; a rapid-entry device, such as a full typewriter keyboard; a light pen, or similar device for providing feedback from the display to the computer and for writing on the displays; audio; and a hard-copy printing device.

In the early years of the digital computer area, major attention of both designers and systems analysts was focused on the arithmetic, control and memory sections of the machine, and little heed was given to the problem of getting into and out of the system. Before long, the limitation of the device was the bottleneck of the I/O for the efficient operation of the system. It is generally recognized today that the total system's effectiveness is determined by the balance between the external communications capability and the workings of the CPU.

Another question that is asked is, "Is there a danger that the computer will impose a rigid and impersonal regime on the classroom, and even replace teachers?" Contrary to popular opinion, the most important potential of computers is to make learning and teaching more an individual affair rather than less so. Students will be less subject to regimentation and moving in lockstep, because computer programs offer highly individualized instruction.

In the work done at Stanford by Dr. Patrick Suppes, for example, he estimated that the brightest student and the slowest student going through his tutorial program in fourth grade mathematics had an overlap of not more than 25% in actual curriculum.

The computer program is neither personal nor impersonal. Whatever else they may be, machines are human thoughts made tangible. The effect and feeling of the program will depend on the skill and perceptivity of those responsible for constructing it. There is little reason to think that computers will ever replace teachers.

The thrust of computer-assisted instruction is to raise the quality of education. In any sort of computer-based instructional system used in the classroom in the near future, teachers will continue dealing with students on an individual basis and doing most of the things they are now doing during most of the school day, with only slight changes.

In looking at education of tomorrow, educational technology will bring changes. For instance, of the four functions of the teacher--that is, interpersonal relationships, environment management, information giving, and clerical duties--the computer will have an effect first on the clerical duties and then on information giving.

Knowledge is multiplying at a geometric rate. It is conceivable that students of the future will be given information on the same basis they are today. Curricula will emphasize teaching--not what the facts are, but how students can gather the facts they need, analyze them and make decisions. The teacher will not necessarily stand in front of a group and lecture, but will be a consultant and a tutor. The library will use the computer and other materials as resources. In the area of information giving, the computer program will handle a greater portion of this function. As a result, the teacher will have more time and opportunity for interpersonal relationships (human-based activity).

In conclusion, we wish to emphasize once again that no one expects that students will spend most of their school hours at consoles hooked up to computers; they will work at consoles no more than 20 to 30 per cent of the time for all uses. The hope is computers will do what Goodlad says, "serve to increase intensity of school as learning environment."

All teachers everywhere recognize the help that books give them in teaching students. The day is coming when computers will receive the same recognition. Teachers will look on computers as a new and powerful tool for helping them to teach their students more effectively.

The new impact, technology in education, is knocking at the schoolhouse door. Its entry is inevitable. It will enter as the result of dedication on the part of American educators that educational goals can be achieved by using this new technology--a technology that puts scientific skills to work by means of various systems of devices led by the most versatile device of them all: the computer--the device which seems to hold the most promise, after some 50 years of search, for individualizing instruction.

CHAPTER XVI TRENDS IN RESEARCH IN THE EDUCATION OF THE HANDICAPPED

M. W. Mueller

The identification and discussion of current trends in research related to the education of handicapped children is a formidable task. Even though data are available, the limitations of these data, and the viewpoint from which they are discussed, warrant first consideration.

One limitation of these observations is that they developed for the most part from data derived from the program of the Division of Research, Bureau of Education for the Handicapped, U. S. Office of Education. The presentation is further limited because the data have been interpreted primarily from the author's own viewpoint, strengthened somewhat perhaps by his vantage point in the Bureau of Education for the Handicapped, but nevertheless from the viewpoint of a single person. Within these limitations, a number of trends in research and related activities in special education are evident.

In discussing trends in special education research, one is implicitly stating that significant dimensions have been identified along which those trends occur. Presumably these dimensions should have specific relevance to the audience to which the discussion is addressed. This paper explores a number of dimensions which may be of more or less interest and relevance to each reader's particular point of view.

The Division of Research of the Bureau of Education for the Handicapped has supported almost $28,000,000 of research over the past 5 years, but this still represents only a portion of the total research effort related to education of the handicapped. One area is of particular interest to the Office of Education, but frequently of fairly limited interest to other educational groups. Many persons tend not to look at the usual categories of exceptionality, which the Office of Education still deals with by law. Nevertheless, it is of special concern in this discussion because the most apparent thing to come out of these data is a rather marked change in the emphasis on various areas of exceptionality. Actually there has been a de-emphasis in the areas of exceptionality.

Division of Research Activities: Categories

Table I provides data on the activities of the Division of Research from 1964 through 1968, broken down into areas of handicap. Handicapped children are defined as those who require special education because they are crippled or otherwise health impaired, severely emotionally disturbed, hearing impaired, mentally retarded, speech impaired, or visually handicapped. The table also includes a number of projects under the heading, "multiple." This does not necessarily mean that the subjects with which a study deals have more than a single handicap. It merely refers to all projects which cannot be clearly identified as involving children who are categorized in one of the specific disability areas. Projects in this area would include those on deaf-blind children, but it would also include projects that do not identify a specific area of handicap, but involve handicapped children more generally. Over the years, the move has been toward greater emphasis on projects that fall into this multiple category.

The field of special education develcped, in many respects, from medicine and psychology. Medical people could readily diagnose a child as mentally retarded, emotionally disturbed, blind, hearing impaired, and so forth. However, the problem of how to educate these children remained, and in many cases it was found that medical and psychological classification systems did not necessarily fit the educational needs of the child. Therefore, the classification system which is imposed by law, as far as federal funding through the Office of Education is concerned, was not quite adequate to the needs of researchers in the field of special education. Over the past 5 years there has been a steady increase in the number of proposals, the amount of money, and the number of projects funded which are not bound to specific areas involving handicapped children. Perhaps

TABLE I

Research and Demonstration Awards by Area of Handicap

Area	Number	%	Amount	%
Crippled or Otherwise Health Impaired	20	7	1,688,589	6
Emotionally Disturbed	31	11	3,476,992	12
Hearing Impaired	46	16	2,315,769	8
Mentaliy Retarded	82	29	4,561,769	16
Speech and Hearing	31	11	3,112,116	11
Visually Handicapped	21	7	1,079,915	4
Multiple	32	11	3,684,742	13
Instructional Material Centers	17	6	5,447,692	20
Research and Development Centers	1	--	2,469,540	9
TOTAL	281		27,837,124	

within a few years we in special education may achieve significant progress and find that we no longer have any classes for mentally retarded or emotionally disturbed children, but only special programs for children with "response deficits" or "response incompatabilities." Such terminology would be more amenable to precise definition, particularly from an educational standpoint. There is an obvious trend to break down the old categorical system, and an increasing amount of research that regards the problems of children in school from some basis other than these somewhat archaic classification systems derived from other disciplines.

The mentally retarded. Over the years, the Division of Research has expended a major portion of its resources to activities related to the education of the retarded. This is a reflection of the severity of the problem of mental retardation in the educational setting. It is the one area of handicap most clearly identified with educational problems. There are blind children who have difficulty with mobility but who do not have problems with education. There are deaf children, a few, who are able to learn very well by visual means. But, since the school's main goal is to teach the child, we must face the reality that few mentally retarded children, by definition, learn well. Therefore, the level of support for projects related to mentally retarded children has been high and relatively stable over the years of our program operation. Nevertheless, the classification, "mentally retarded," tends to obscure one problem. That is, that the mentally retarded are not a homogenous group. Mentally retarded children are ranked as mild, moderate, severe or profound, or are classified as educable, trainable, or custodial, according to different viewpoints. Roughly 90% of the research supported by the Office of Education has been related to mildly retarded children. This trend is gradually reversing. The level of support of research in mental retardation has remained stable, but there is increasing emphasis on trying to provide support for research on educational problems of retarded children with moderate, severe, or profound learning deficits.

The emotionally disturbed child. In the area of emotional disturbance, the trend is downward in level of support and number of projects. The principal reason for this is that the Bureau of Education for the Handicapped has become more and more sensitive to the support of research that is strictly related to education. A great many children who have problems that are identified as emotional disturbance of one degree or another actually do not have major academic problems. We are inclined to limit our support to activities which relate to the child whose particular disturbance hinders his educability.

The hearing impaired child. Research on hearing impaired children has been fairly steady with a slight tendency to increase in recent years, primarily on one basis. In education generally, there has been a strong indication to give greater attention to the children at very young age levels. The field of deaf education for many years has been sensitive to the problem of educating children as young as 18 months of age, and, in many cases, various pre-academic or pre-educational efforts have been undertaken even below this age. Therefore, as we have become more sensitive to the needs of preschool children and augment our knowledge about the educational process in the preschooler, the field of hearing impairment has received a slightly larger percentage of our resources.

Speech and hearing defect. Speech and hearing, primarily projects in speech pathology, have also shown a slightly greater than average increase over the last few years. The reason for this is more difficult to discern than some of the other trends in research in various areas of exceptionality. It does seem that there is a greater interest in research on speech pathology, speech correction, and various other projects related to the speech process in public school systems. In the past, research into problems of speech and hearing were largely limited to clinics associated with hospitals. The word "pathology" even gives some people the feeling of a medical rather than an educational orientation. In recent years, there has been an increased interest in speech in the school setting, and as a result there has been a greater emphasis in this area in research activities as well.

Crippled and other health impairments. The last two areas, visually handicapped, crippled, and other health impaired children, have received relatively limited support. The majority of research in both these areas tends, as it used to in speech and as it certainly is in some aspects of emotional disturbance, to be more medical than educational in nature.

Consequently, we in the Office of Education have not devoted a major portion of our resources to these areas. In the field of crippled and other health impaired children, there is a tendency, even in general fields like special education, to pay somewhat less attention than previously. The crippled child with normal intelligence is more and more being regarded as a problem requiring some slight environmental adaptations within a fairly general educational program rather than a special education problem. This is not so true of the visually impaired child and the only apparent explanation of the level of support which we have devoted here is that there are very few researchers expressing interest in educational problems related to vision.

Overall, these trends boil down essentially to the one factor that was noted in connection with multiple handicaps. Although there probably will continue to be a great deal of categorical research, more and more the problems that we face in public schools or institutional settings, the various educational settings dealing with handicapped children, do not necessarily fit easily into the old categorical systems to which we have been accustomed.

Second Dimension: Research Strategies

The second major dimension which appears to have some relevance is the area of research strategies utilized in the field of special education. The majority of educational research conducted in the field of special education has been tied to what one might call an experimental psychology design--identifying experimental and control groups, testing those groups, treating the experimental groups, testing both groups again, and analyzing the differences between them. Although this design is appropriate to many problems, it tends to discount, or, at least, inadequately control the antecedent condition, the comparability of those groups, and the individual differences within the groups.

Another type of research supported by our program that has been conducted generally in the field of education is correlational research. Again, this research methodology leaves a great deal to be desired if you want to know something about an individual child. It is effective for determining how closely related intelligence and achievement are in in a mentally retarded or in an emotionally disturbed population, but it doesn't help a great deal with planning for the individual child. In addition, there is a tendency to conclude that because there is a close relationship between intelligence and academic achievement in a mentally retarded population, the lack of academic achievement is due to a lack of intelligence. This is a fairly logical conclusion, but there is no certainty that the relationships observed are necessarily cause and effect. In fact, there has been some reasonably good evidence that certain features of performance reveal very little difference between the retarded and the normal population! In addition, data from correlational studies are in many cases inadequate in providing solid ideas of how we can now teach a child in order to reduce the correlation between intelligence and academic learning.

Perhaps as much as half the research being supported by the Division of Research falls into one of these two classifications. This is mainly because the field is still very much oriented in this direction. Research personnel know how to use these experimental designs and to use them well. Therefore, we are continuing to support research which must be interpreted with caution since it is based on group data. This does not mean that there isn't a use for such research, but certainly it can't be the whole story and more and more it becomes a smaller part of the whole story.

Behavior analysis and modification. We are giving support to projects in the area of the functional analysis and modification of educational behavior. This is a strategy of research which is being used increasingly by those researchers interested in the education of the handicapped. In addition, research strategies that have been developed in the fields of sociology, cultural anthropology, and other disciplines are at least beginning to have an impact on educational research.

Research Content: The Third Dimension

Another general dimension where some changes are apparent is that of research content. If we analyze the content of the research which we have supported over the last few years, we find it can be broken down into about six types of activity. These include characteristics of handicapped children, assessment, special programming, curriculum and materials, administration, and teacher training.

Characteristics of handicapped children. As we review the early years of our program, and, indeed, as we look at the broader picture of special education research in the United States, we note that much of the research activities was limited to studies of the characteristics of children. A recent analysis of projects supported by the Division of Research indicated that more than half the projects dealt with characteristics of handicapped students. Unfortunately, in many cases these characteristics had not been studied in such a way that they were directly relevant to what might be done for the child in an educational setting. The direction now is for a decrease in this emphasis in our particular program, and perhaps in the broader field of research in special education. To the extent that such studies will continue to be supported by federal funds through the Office of Education, they will at least be limited to studies of those characteristics that can lead to some direct use by teachers of handicapped children and related personnel.

Behavioral repertoire assessment. A second major area of content is assessment of the handicapped. In the past, a great many projects have been involved with determining how to find out more about handicapped children. Overall, there seems to be a decreasing trend here.

Assessment, like studies of characteristics, is of little value except as it provides information useful in teaching the handicapped child. Two major directions show promise of providing assessment techniques and tools to accomplish this. One of these is research on assessment relating directly to how a child learns in a classroom. Instead of providing the teacher with an IQ score, we would much rather provide a profile of abilities that are directly relevant to educational performance. A number of additional tests in the same vein, trying to provide information about a child that has immediate relevance to a school situation, are being researched in the field today.

In the past, most of the instruments which were available for assessment measured the products of intelligence. They measured the intelligence by determining what the child had learned in the past. At present, there is a rather significant move toward measuring intelligence by what the child can do. In behavioral terminology, this would be analogous to developing an intelligence test based specifically on baseline response rates.

People who are doing research in the area call it fluid versus crystallized intelligence, or the measurement of intelligence by process rather than product. Whatever the terminology, the aim is similar. If you want to be able to predict how a child will learn over a period of years, rather than ask him questions about how he learned over a certain extent of time, it would be better to sit the child right across the table from you and see how long it takes him to learn something. A few tests are now available which are based on this assumption, but work is still needed. This is probably one of the more significant trends in assessment research.

Special programming. Special programming research efforts have been fairly important in the past, but perhaps the most important information has already been revealed to us: that special programming must go beyond simply segregating children, or assigning a teacher fewer children. We have demonstrated fairly well that special classes aren't enough in special education. Further efforts must be directed toward programming that meets the specific educational needs of the handicapped.

Curriculum and materials. The next area, curriculum and materials, represents one of the major increasing trends that we see. Those of us who deal with handicapped children do not have enough explicit information regarding curriculum that is particularly appropriate for the handicapped child. Nor do we have enough materials to implement curricula that are especially suited to a handicapped child. Neither do we have information on the methodology of communicating through materials to reach curriculum goals. This is an area, a very broad area, where research is definitely on the increase.

Administration and teacher training. Two other areas which have had very little attention in the past appear to be coming developments. These are administration and teacher training. In our particular program in the Bureau of Education for the Handicapped, our support of research activities has been limited almost exclusively to activities dealing with how to treat children or what children are like. In the future, we will be giving much greater attention to the total educational environment, the organization and administration of special programs for handicapped children. We are also going to be devoting much more attention than we have given in the past to research on how to more effectively train teachers. With the exception of a few projects, we have not yet delved in this this particular area.

Applied Versus Basic Research: Fourth Dimension

Related to research content is the area of applied versus basic research. Again this is a dimension in which some changes are apparent. The Bureau of Education for the Handicapped is mandated by Congress to support applied research. Nevertheless, five years ago when this program came into existence, there was a rather strong tendency for people to use a slightly different breakdown between what is basic and what is applied research than we do today.

Previously, most people characterized basic research as research that was information oriented rather than problem oriented. It is sufficiently controlled so that there can be a great deal of confidence in the validity of the outcomes, even if those outcomes may not be terribly relevant to immediate practice in a classroom situation. Applied research has been characterized as problem oriented, atheoretical, and loosly controlled. These viewpoints have characterized applied research as poor research, even though the results in many cases were valuable to the teacher. Over the past five years, we have been seeing more and more good applied research--research that is tightly controlled and that you can have a great deal of confidence in--with results that are directly applicable to teaching children.

Five years ago we also supported a good deal of basic research, because there was not enough good applied research to utilize our resources well. Today, more and more good applied research projects are being supported and basic research is being turned back to the National Institute of Health, National Science Foundation, and to similar organizations whose support is not limited to applied research.

Research Settings: Fifth Dimension

A fifth dimension is concerned with the settings in which research is conducted. The diversity of sources of federal funds and the criteria by which different agencies work make it very likely that any given agency will have a rather discrepant sort of breakdown for various types of institutions. Table II indicates that over the history of our program about two-thirds of our resources have gone into college and university settings. This has been fairly consistent during the five years of the program. Schools, and this includes public schools, private schools, institutional schools, residential schools, and other school situations, have received about 15% of the projects and about 8% of the money distributed by the Division of Research. Over the five years of our program, the trend reflects a decrease in support of research in public school settings.

One might have expected that public school systems were becoming more interested in research--more interested in finding out whether the programs they were conducting were really doing the jobs intended. On looking further into some of the specific proposals and projects which make up these data, however, it is possible to discern what is happening. During the early years of our program, the public school had almost nothing available to it in the way of federal support. As a result, many projects submitted the first year of our operation were actually to provide services but they were presented as research or demonstration projects in order to obtain support. Now that there has been a very great increase in the amount of federal money available to public schools under the Elementary and Secondary Education Acts, the need for this minor deception no longer exists. As a result, there is a decrease in the amount of support given to research in public school settings since funding from other sources is available to public school people.

State education agencies receive a small portion of our funds as do clinics and hospitals, professional organizations, specific research organizations, but none of their data reflect a sufficient number of projects to discern any trend. State departments of education are certainly very interested in doing research, but, like the situation in public schools, there are financial resources available to support this research. Clinics and hospitals have also maintained a relatively stable rate, perhaps a slight decrease. Again, when our program first came into operation, speech clinics, psychiatric clinics, and other groups considered federal funds a major resource. As the parameters of the research which

TABLE II

Research and Demonstration Awards by Type of Institution

Type of Institution	Number	%	Amount	%
Colleges and Universities	185	66	18,651,834	67
Schools	41	15	2,339,993	8
State Departments	11	4	1,512,673	5
Clinics and Hospitals	13	5	1,925,752	7
Professional Organizations	11	4	1,250,941	4
Research Organizations	11	4	1,388,071	5
Other	9	3	767,860	3
TOTAL	281		27,837,124	

we will support have become more clearly defined as being strictly educational in nature, the clinics and hospitals have reduced their interest somewhat, although there are enough handicapped children in these settings so that the level of interest from such institutions remains relatively stable.

Observations of these trends and broader considerations in the field of special education have led the Bureau of Education for the Handicapped to clarify its role in the support of educational research related to the handicapped. On the basis of these considerations, the Bureau's Division of Research has developed a more specific delineation of its objectives and the dimensions of its program.

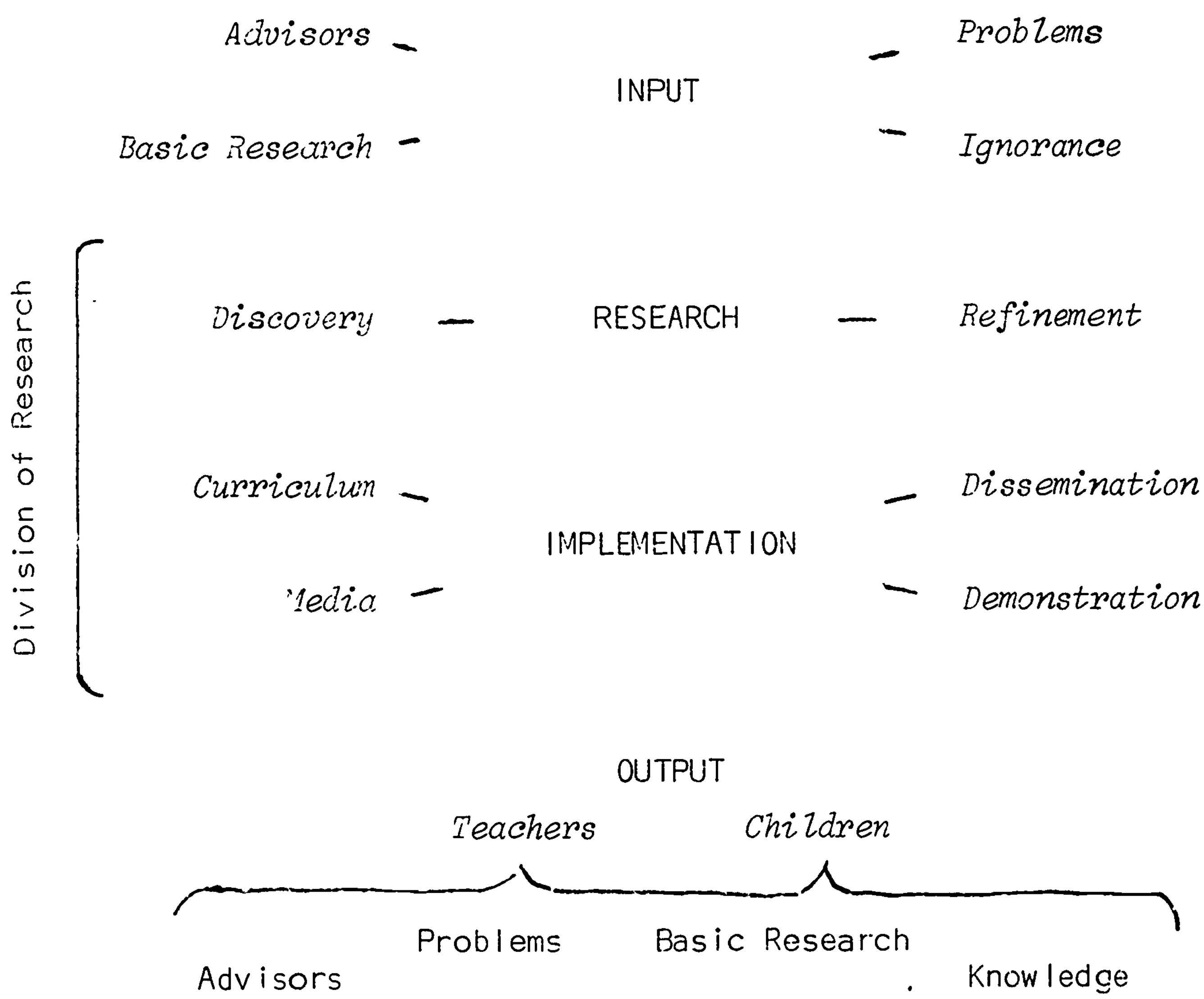
The single purpose is to find ways to provide handicapped children with an education commensurate with the complexities of living in tomorrow's society--we know we can do better than we are doing today. Our job is to find ways to make this possible. To be more precise, our goal is to improve educational opportunities for handicapped children through the support of applied research and related activities. Therefore we must be more specific about this overall objective, and refine our knowledge by obtaining more precise information relating directly to educational programming for the handicapped. Related activities include efforts to implement the information developed in the research program. Besides supporting activities related to these major program goals, additional resources are devoted to certain activities designed to maintain the viability of the total program. The generalized model describing the program of the Division of Research is presented in Table III. As indicated there, the program is involved mainly with research and implementation. Program maintenance activities have some involvement with input and output, but bear little direct relation to the generalized model within which we operate.

Applied research. The first major segment of the program is the support of applied research activities. This is an extremely broad program with responsibility for the development, refinement, and manipulation of information relevant to all aspects of education for handicapped children.

Implementation. The second major responsibility, program implementation, is designed to put information derived from the research program

TABLE III

GENERALIZED APPLIED RESEARCH MODEL



into practice to improve educational opportunities for the handicapped. The specific activities supported under this program include: dissemination, demonstration, curriculum, and media projects.

In many instances information needs only to be channeled into the proper areas to be useful in the improvement of education for the handicapped. One responsibility of the implementation program within the Division is to provide support for activities that will ensure that such information is indeed disseminated efficiently and to all potential consumers.

In other cases, dissemination of information does not represent the necessary and sufficient conditions for implementation of research information. One additional step which is often necessary is the demonstration of that information. In such cases it is expected that only the information that is soundly based in research will be demonstrated, that it will be demonstrated under replicable conditions facilitating generality of the data, and that such demonstration will be readily accessible to all potential consumers.

In still other cases, the information derived from the research program cannot be adequately implemented without the development of an appropriate vehicle to carry that information. The typical vehicle for implementation of educational research is the curriculum. To assure adequate implementation of all information, curriculum development activities designed to provide appropriate vehicles to implement all research information are supported.

In still other cases information cannot be implemented through simple dissemination, demonstration, or even with the development of new vehicles. It may be necessary to support the development of new tools to implement this information. For this reason, activities are supported that are designed to develop these tools and the media required to complete the implementation process.

Summary

In summary, trends have been observed in connection with the programs relating to areas of exceptionality, research strategies, research content, applied versus basic research, and the setting in which research is being conducted. In terms of areas of handicap, there is a good deal less emphasis on individual handicapping conditions and more on educationally handicapping behavior. regardless of the underlying conditions. In terms of research strategies, the trend veers from what was developed in psychology, medicine, and other fields, toward a methodology that can be adapted specifically to educational problems. In terms of content we seem to be moving away from the study of the characteristics of pupils or even their assessment, toward more research on curriculum, materials and methods, and how these interact.

Finally, there is some hint at least that we are going to be increasing our attention to administration, organization, the physical environment of the child, and the training of his teachers from a research point of view. The move toward research which is more applied by nature, and the fairly stable rate observed in terms of research settings, represent fairly clear trends, but are probably quite restricted to the program of the Division of Research.

These are some dimensions which may have some validity for us as educators, and where available, the data have been amenable to discerning some trends. From this information, a rationale for describing and limiting the program of the Division of Research has been derived.