DOCUMENT RESUME

ED 033 842 24 RE 002 326

AUTHOR Silverberg, Norman; And Others
TITLE Decision Models in Remedial Reading: A Pilot Project. Final Report.
INSTITUTION Kenny Rehabilitation Inst., Minneapolis, Minn.
Spons Agency Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No BR-8-F-119
Pub Date 30 Sep 69
Grant OEG-6-9-008119-0049
Note 41p.

EDRS Price EDRS Price MF-$0.25 HC-$2.15
Descriptors Auditory Visual Tests, Corrective Reading, Elementary Grades, *Individual Characteristics, Individual Differences, Kinesthetic Methods, *Methods Research, Pilot Projects, *Predictive Ability (Testing), Predictive Measurement, Reading Research, Reading Tests, *Remedial Reading, *Simulation, Visual Perception

Abstract

In developing a model to utilize simulation techniques for assigning remedial reading to children, 100 third graders in a St. Paul, Minnesota, school district were placed in 16 remedial classes at 12 schools under the supervision of seven experienced remedial reading teachers. All of the subjects had scored 90 or above on a group IQ test and were designated as reading one grade below placement. The classes were divided for 8 months of training into three groups, each taught by different remedial methods: one visual, one kinesthetic, and one auditory-phonic approach. Pretest batteries to dermine skills in perceptual, perceptual-motor, and verbal areas and three reading tests were administered. The average of two reading tests was used as a post-test criterion. Prediction equations in all three methods were applied to the post-test scores in order to compare simulated reactions to the two methods not used with a pupil to his actual reaction to the method used. More than half the total sample predictively would have achieved a tested reading level of from 7 months to 1 year higher had they been placed in methods other than the one to which they were assigned. This suggested that children's reactions to a remedial treatment can be predicted by their individual characteristics. Tables and references are included. (BT)

BR 8-F-119
PA 24
OE/BR

FINAL REPORT

Project No. 8-F-119

Grant No. OEG-6-9-008119-0049 (010)

DECISION MODELS IN REMEDIAL READING: A PILOT PROJECT

Norman Silberberg, Ph.D.
Iver Iversen, M.S.
Jean Goins, Ph.D.

Kenny Rehabilitation Institute
1800 Chicago Avenue South
Minneapolis, Minnesota

September 30, 1969

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT

Project No. 8-F-119

Grant No. OEG-6-9-008119-0049 (010)

DECISION MODELS IN REMEDIAL READING: A PILOT PROJECT

Norman Silberberg, Ph.D.
Iver Iversen, M.S.
Jean Goins, Ph.D.

Kenny Rehabilitation Institute
Minneapolis, Minnesota

September 30, 1969

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

ACKNOWLEDGEMENTS

This project could not have been completed without the work of Mrs. Arlene Gillett and Miss Marylyn Hoglund of the research staff of Kenny Rehabilitation Institute. In addition, the cooperation and dedication of the following remedial teachers in the North St. Paul-Maplewood School District are acknowledged:

Mrs. Beulah Carlson
Miss Beth Fangman
Mrs. Marjorie Graber
Mr. Loren Johnson
Mrs. Claire Obsatz
Mrs. Georgia Prakash
Miss Lois Toedter

TABLE OF CONTENTS

SUMMARY

Remedial reading is an expensive and widely used educational technique. The remedial teacher is forced to make a priori decisions as to the method and materials she will use for a specific child. Educational research has provided no concrete guidelines for this decision. This study has addressed itself to the problem of matching individual children with the remedial techniques having the highest likelihood of being successful. To this end, three replicable remedial techniques, including a visual, a kinesthetic, and an auditory-phonic method, were investigated to determine those learner characteristics associated with positive and negative responses to each technique. It was found that, through the use of simulation techniques, operational relationships could be established between a child's characteristics and his response to the three remedial methods. This model appears promising as a means of allocating children to a remedial method appropriate for him, thus maximizing the impact of remedial reading.

Chapter I

A DILEMMA IN REMEDIATION

SUMMARY

Remedial reading is an expensive and widely used educational technique. The remedial teacher is forced to make *a priori* decisions as to the method and materials she will use for a specific child. Educational research has provided no concrete guidelines for this decision. The project described herein has addressed itself to the problem of matching children with remedial techniques in such a way that each child would be assigned a technique most likely to be successful in that child's remediation.

As the awareness of the effects of learning disabilities increases, more and more programs are being developed in an attempt to remediate the condition. New alphabets, programmed approaches to learning, and individual or small group tutoring are among the many techniques attempted with children with learning difficulties. It is reported that there are 12,415 primary grade school children in public and parochial schools undergoing formalized supplementary reading and language instruction in Minnesota alone, these programs being funded under Title I Projects of Public Law 80-10*. Supplementing these school programs, many churches and neighborhood organizations offer their own afternoon programs with the proclaimed goal of improving the study skills of "deprived" children. In addition, many children are referred to child guidance centers because of "emotional blocks" to their learning.

The child with a reading problem invariably faces the frustration and anxiety of perceiving himself, and being perceived by others, as a failure. As such, he is subjected to a remedial experience which may or may not be successful. If it is unsuccessful, it merely extends the time that the child is involved in this unhappy situation. In Minnesota, only one-half of the primary grade children undergoing remedial reading are perceived by the local school administrators as having achieved "substantial progress" as a result of this extra work.

*Personal communication from Mr. C. Bezenson, assistant administrator in Title I Evaluation, Minnesota Dept. of Education, 1966.

Such painful results have long been evident to those working in the field of remediation, and much research has been undertaken in the hope of achieving greater success by using one remedial method or another. However, in most research in which a specific remedial method is studied, the only inference on completion of the study can be: the method worked adequately (according to some "class average" criterion of adequacy) or it did not. The amount of improvement shown, however, is actually distributed over a range -- with some children showing an encouraging amount of improvement, some children evidencing no success, and some children falling between these two groups. We do not know whether a specific child might have improved more or less with exposure to some other method.

In applying the results of research to classrooms, decisions are often based on a comparison of mean score changes in reading level on one or more standardized reading tests. Unfortunately, statistically significant differences between groups may be so small as to be of little practical use to the teacher. If several remedial methods are available, the teacher or reading consultant must make a selection based on a priori predictions of children's responses to the available remedial materials, these predictions based on behaviors and/or test scores. Typically, "In the treatment of the scores of children on a variety of tests as an aggregate, the individual learning ability of the specific child may be masked," (Haring and Ridgeway, 1967).

As a result, there are a number of instruments available for diagnosing children, there are a number of treatment methods for remediation, and there are relationships established between results of diagnosis and results of treatment on the average -- but no empirical relationships have been established between diagnoses, and for a number of treatments. Although certain relationships have been intuitively postulated (e.g., perceptual training will aid children with visual-perceptual problems, or conversely, children with visual-perceptual problems should receive auditory training to compensate for this), the common characteristics of children who respond to a particular remedial method have not been identified.

The question then arises: Might there be a more efficient way of deciding whether to present the child with remediation, and into what type of remedial class might he go? If there were such models for decision making, we could then decide whether a child would benefit most from a particular remedial method, would benefit equally under any method, or would benefit from none of the available reading methods.

It would seem reasonable that different children learn best through different approaches. It was therefore hypothesized that ways can be formulated which delineate treatments appropriate for subgroups of children possessing certain common characteristics (some of these common characteristics possibly being diagnostic).

This study offers both a test of this hypothesis and a first step in the direction of placing the child in the remedial method appropriate to him. Since there is a plethora of theories concerning reading retardation, it would seem more efficient to define children operationally according to their response to a treatment method rather than to await validation of each of the many theories underlying reading problems. This approach is common, e.g., dexedrine is commonly administered to hyperactive children despite the fact that the reasons for this drug's effectiveness are still in the investigative stage.

The mere fact that there are so many remedial methods in popular use suggests that most well-thought-out theories on reading disability receive a measure of acceptance from some groups of reading specialists. Often, textual series used in schools draw on several theories. The proposed study limits the remedial methods to textual series. This approach is consistent with the broad concept that reading retardation may be due to improper training methods.

In this study, interactions between a large number of diagnostic variables and three methods of remedial reading were investigated. An attempt was made to develop operational methods of predicting an individual child's reaction to one of the three methods. If this approach proves feasible in practice, some of the burden of prediction will be removed from the teacher who can then devote her energies to teaching rather than using important instructional time in short-term trials of several types of material in the hope of gaining a clue as to which material is most appropriate for a particular pupil.

Further research into the relationships between student characteristics and the many different types of remedial approaches in current use would be necessary in order to generalize the approach developed in this study. However, a generalized application of this approach could yield an increment in student achievement well beyond current expectations.

Therefore, this study was designed to provide a model which could be replicated in other educational programs. Its purpose was to develop a basis for quantitative decisions in the allocation of individual children into one of several alternative remedial teaching methods available to such programs.

Chapter II

POPULATION AND SUBJECTS

SUMMARY

This study involved all third grade children in sixteen remedial classes in twelve schools in one suburban school district. Inclusion depended upon scoring 90 or above on an IQ test and being one grade retarded in reading.

It was observed that children who develop difficulties in adjusting to their environment most often have histories of school failure. It seems wise, therefore, to be concerned with the prevention of such problems through early detection and remediation. With this preventive goal in mind, the study dealt with children in third grade, since a child cannot truly be classified as a reading problem until that time. By and large, second grade reading programs offer beginning reading to children who were unsuccessful at the first grade level. If a child is again unsuccessful in learning to read during second grade, he is improperly prepared for third grade level work which presumes an adequate skill in the reading process.

In the North St. Paul-Maplewood school district, reading retardation is defined in third grade as normal IQ (90 or above on group IQ tests) and one year retarded in reading. It is recognized that group tests are designed to function as screening devices and that an IQ of 90 on such a test may not be a valid indicator of the outcome on an individually administered instrument (Fitzgerald, 1960). Nevertheless, group test scores are typically all that are available to the third grade teacher; the shortage of school psychologists obviates the possibility of individual IQ tests for all referred students as a condition for acceptance into a remedial program. The selection of the 90 IQ cutting point was arbitrary; however, Lytton (1967) demonstrated that high IQ students benefited more from remedial instruction than low IQ students. For this reason, the criteria for inclusion in the experiment seemed feasible from a practical point of view.

The experimental subjects consisted of children in the North St. Paul-Maplewood Schools who were entering third grade in either a public or parochial school for the 1968-69 school year, who had not yet been in formal remedial programs and were referred by their classroom teachers for remedial reading. Efforts were made to encourage referrals on all children who read poorly and who scored 90 or above on a Lorge-Thorndike Intelligence Test administered either in second grade or prior to beginning the remedial program in third grade. The classroom teachers were instructed to refer children who were reading one or more grades below grade placement according to their evaluations.

The subject pool appears quite representative of a population of children with reading problems, including urban, suburban, and farm children, representative of most racial and ethnic groups, and representative of families with incomes over the entire range represented in a public and parochial school population. Table I describes certain demographic characteristics of the population.

Table I

Frequency Distribution of Selected Demographic Variables

Parents' Marital Status:	
Married	92
Separated	0
Divorced	6
Widowed	2
No Data	2
Father's Occupational Status:	
Craftsman-Foreman	28
Operatives	16
Managers, Officials, Proprietors	9
Professional	8
Clerical	7
Laborer	7
Sales	5
Service Workers	4
No Data	18
Mother's Occupational Status:	
Housewife	89
Clerical	4
Operatives	1
Service Workers	1
No Data	7

Occupations were classified according to categories established by the U. S. Census Bureau in the 1960 Census Population: Index of Occupations and Industries. The most frequent occupational classification of fathers was that of Craftsman-Foreman, with 28 of the 84 identified falling in that category. The next most frequent classification was Operative, with 16. Eight fathers were classified as Professionals, and 9 as Managers or Proprietors. The remainder fell in the Clerical, Sales, Service Worker, or Laborer classifications. All but 8 of the mothers were classified as Housewives. Only 8 of the children came from homes where the parents were separated, divorced, or widowed and therefore lived with only one parent.

A total of 102 students were included in sixteen classes in eleven public schools and one parochial school. There were two boys for every girl selected, a smaller ratio than expected (N. Silberberg and Feldt, 1968). This sample represents approximately 10 percent of the total third grade population served by this remedial program, which is consistent with the expected percentage of children to be included in a remedial program (N. Silberberg, Iversen, and M. Silberberg, 1969).

No control subjects were involved in the study. It would have been interesting, but not critical to this experiment, to determine the progress of children receiving no remedial help randomly selected from the referred group. However, the advantages would be outweighed by the moral consideration of withholding treatment from a child in need, as well as by the difficulties encountered in teacher morale. It is expected that, were children randomly withheld from treatment, school staff cooperation would have been significantly reduced. Teachers are typically too service-oriented to accept administrative procedures which they perceive as harmful to even a single child.

Chapter III

REMEDIAL METHODS UTILIZED

SUMMARY

One of three commonly used remedial approaches was randomly assigned to each remedial class. All of these techniques (kinesthetic, visual-sight, and auditory-phonic) are based on an appeal to a specific and distinct sensory modality.

Three remedial approaches were involved in this research:

1. Visual-Sight Approach
2. Auditory-Phonic Approach
3. Kinesthetic Approach

The three techniques were selected as the optimum number both on the basis of their representativeness of remedial techniques in general and the statistical consideration of including an adequate number of pupils in each sample.

Considering the large number of remedial techniques available, the selection of three will certainly exclude some methods of merit. Several criteria were invoked in this selection. As pointed out by Bliesmer and Yarborough (1965), there is no research evidence on the superiority of one or more of these techniques in comparison with other techniques. They are all popular, if popularity can be rated by the number of times they are mentioned in professional literature or at professional meetings. The following criteria were utilized in selection:

(1) The methods were selected so as to represent the main categories of remedial methods, each method concentrating on a specific sense modality. The three methods chosen represent a visual-sight approach, kinesthetic approach, and an auditory-phonic approach. Certainly none of the methods are "pure" in their exclusion of aspects of the other approaches, but the methods are typically viewed as distinct by reading specialists. The selection assumes that if a child would respond to a certain extent, let us say, to the selected auditory-phonic approach, he will respond in a fairly similar way to a different auditory-phonic approach, as suggested by Bliesmer and Yarborough's (1965) results.

(2) The methods were selected for use in the remedial classrooms. This eliminated the use of some promising methods which are administered individually (e.g., Heckelman, 1962) and the use of developmental reading methods meant for use in the regular classroom. It should be noted, however, that the subjects of this study had already encountered failure with regular classroom techniques.

(3) The methods selected require no special training for the teacher. In order to maximize the utility of this research to other school districts, the techniques were selected so that they could be taught by the school's regular personnel. This eliminated the use of some motor training techniques which are currently popular.

(4) Each remedial technique selected was, in the opinion of the Project Supervisor, acceptable to the remedial teachers as a teaching method.

A. Visual-Sight Approach -- The preliminary training in this program was the Frostig Program for the Development of Visual Perception. A diagnostic test supplied with the materials was used to determine areas of weakness; subsequently, the training focused on areas identified by the test. The visual training continued as long as the tasks presented challenged the pupil.

Concurrent with visual training, aid in using contextual clues to identify unknown words was introduced. Oral exercises and also exercises using a cloze technique (leaving word space blank) were used to develop ability to use context as an aid in word recognition.

Tachistoscopic presentation of words in which the teacher initially identifies the words, and subsequently the pupils mark the correct word on a line of similar words, and eventually the pupils call out the word when it is presented, followed the Frostig training sequence.

B. Auditory-Phonic Approach -- The basic approach was a "synthetic" presentation of phonic elements of the language. Letter names and sounds were learned in isolation initially and gradually combined and blended into longer word forms. The Gillingham-Stillman manual was used by the teachers as a guide for the systematic, sequential presentation of the sound elements to be learned. When necessary, additional practice material for sounding and blending words was derived from the Hegge-Kirk reading exercises.

C. Kinesthetic Approach -- The technique developed by Grace Fernald was used. Initially, the pupils were encouraged to write stories about topics of their own choosing. The pupil could ask the teacher for any word he wished to learn. The word was written in large script with crayon on a strip of paper. The child then traced the word with his finger(s), saying each part of the word as he traced it, until he felt he knew it. The word then was covered and written from memory (never by looking at the copy). This process was repeated as often as necessary, the word always being written as a whole. If errors occur, words were crossed out and begun again.

When a story was completed, the teacher printed or typed it and made it available for reading by the pupil as often as he wished. Each pupil had an alphabetical file box in which to store his word strip for future reference. Copying words was not allowed; the file was for relearning words by tracing if the word had been forgotten.

One weakness in the comparison of these methods was the fact that not only were the teaching methods different, but the materials used in conjunction with these methods were different as well. Each series of texts differs in format, vocabulary, storyline, etc., with the resultant possibility of differential interest and motivation on the part of the student. It was hoped to be able to adapt one textual series to use for practice with narrative materials for all three methods. This, however, was beyond the scope of a pilot study.

Chapter IV

METHOD

Research Design

SUMMARY

One hundred and two students participated in this program, with approximately equal thirds randomly assigned to each of the three methods. Seven remedial teachers were assigned to the three methods in such a way as to attempt to equalize the teacher factor.

Sixteen remedial classes were involved in the study, yielding a total population of 102 students. There was no reason to expect that the subsets of subjects included under each remedial method were dissimilar.

The teachers were instructed to teach the children as normally as possible, the only restriction being the use of the educational materials assigned to them. This is not an unusual practice, either in terms of the teachers or the students. Many school systems in the area have "specialized" in certain techniques and the remedial teachers are expected to adhere to the guidelines. For example, one large suburb in the Twin City Area is using the Gillingham-Stillman approach in special classes for educable mentally retarded, and the McGraw-Hill programmed materials are being used in regular and supplementary classrooms on both the elementary and secondary levels in North St. Paul.

The problem of teacher-method interaction is an important one but beyond the scope of this pilot project. This problem could be partially avoided by a random allocation of method to class, which would mean that most teachers would be teaching a different method in each of their two classes. Although theoretically appropriate, this was not practically feasible because one could be almost certain that each teacher would select certain aspects of the method taught in one class and apply it to her other class, thus reducing the "purity" of each method. Each teacher found it more efficient to duplicate a single method in two classes. Therefore, investigation of teacher-method interaction would be more appropriate in larger scale follow-up studies.

To minimize teacher effects, the Project Supervisor ranked the seven teachers from "best" to "worst" in terms of experience during the 1967-1968 school year. The criterion of "best" was observed ability to interact in remedial classes. No teacher was inexperienced. Teachers were then allocated to method in such a way that the mean rank for each method was approximately equal.

Data Collection

SUMMARY

Each child was measured on many quantifiable psychological, educational, and demographic variables. The variables were selected because of their theoretical, already experimentally demonstrated, or logical relationship to school functioning.

Each child, following selection for inclusion in the experiment (based on Lorge-Thorndike IQ and informal reading inventory), was administered the psychometric test battery by a psychometrist during the first five weeks of the semester. This test battery was designed to test many of the abilities which have been found to be related to reading achievement. These tests were arbitrarily selected, based on their experimentally demonstrated relationships with reading problems and on the clinical experience of the Project Director and Project Supervisor. The battery included:

Gilmore Oral Reading Test (Reading paragraphs aloud)
Wide Range Achievement Test (Reading words in isolation)
Bender-Gestalt Test (Test of visual perception)
Wepman Auditory Discrimination Test
Peabody Picture Vocabulary Test (Test of recognition vocabulary)
Benton Visual Retention Test (Error and Correct scores)

During the first week of remedial classes, the remedial teacher administered certain educational tests commonly used in the third grade, including group administrations of the Draw-A-Man Test, the Botel Reading Inventory, and the Gray Oral Reading Test. In addition, Metropolitan Achievement Tests were available from second grade testing.

During the remedial semester each child was administered a Wechsler Intelligence Scale for Children by a psychometrist. Since this instrument yields standard score equivalents for given chronological age ranges, there was no necessity to have it administered at a specific time during the semester.

The child's sex and chronological age at the beginning of remedial instruction were also included in the data pool.

The results of this testing for the total sample can be found in Table II. It can be seen that some of the children included in the remedial program had, in fact, tested reading levels higher than that suspected by their second grade teacher or their third grade remedial teacher. Some aspects of the problem of defining reading level have been discussed elsewhere (N. Silberberg, Iversen, and M. Silberberg, 1969), but the problem is far from settled. Evidently, there is considerable discrepancy between tests of reading, and between any test and the child's "functional level" as observed by the teacher. The lack of quantitative definitions of reading difficulties compounds this problem significantly.

Mean scores on selected pre-test variables are included in Table III. It can be seen that prior to remedial treatment, few significant differences existed between the three groups. Most differences were of little practical importance save one: the kinesthetic group scored higher on reading tests than did the other group.

Table II

Means and Standard Deviations of Test Variables

Test	Mean	S.D.
Chronological Age (Months)	100.5	5.6
Number of Siblings	3.5	2.3
Metropolitan Achievement Test		
Word Knowledge (2nd Grade)	2.2	.3
Word Discrimination (2nd Grade)	2.4	.5
Reading (2nd Grade)	2.1	.5
Lorge-Thorndike IQ	100.55	8.96
Draw-A-Man Standard Score	94.04	13.58
Pre-test Gray Oral Reading Test,		
Grade Equivalent	1.98	.63
Botel Reading Inventory		
Word Recognition	4.00*	.23
Phonics, Raw Score	41.94	12.27
Listening Comprehension, Raw Score	47.67	9.39
Peabody Picture Vocabulary Test		
IQ	104.6	14.0
Raw Score	70.2	8.1
Gilmore Oral Reading Test		
Accuracy, Grade Equivalent	1.9	.7
Comprehension, Grade Equivalent	2.4	.9
Speed, WPM	64.2	17.7
Wide Range Achievement Test		
Reading, Grade Equivalent	2.3	.4
Reading, Standard Score	87.8	5.2
Spelling, Grade Equivalent	2.2	.4
Spelling, Standard Score	86.9	5.5
Bender-Gestalt, Raw Score	2.8	2.4
Benton Visual Retention Test		
Error	9.2	3.2
Correct	4.9	1.5
Wepman Auditory Discrimination Test,		
Raw Score	3.9	3.1
Wechsler Intelligence Scale for Children		
Information	9.70	2.13
Comprehension	8.95	2.50
Arithmetic	9.55	2.38
Similarities	10.83	2.61
Vocabulary	9.74	2.78
Digit Span	9.36	2.39
Picture Completion	10.37	2.96
Paired Associates	11.11	2.63
Block Design	10.78	2.60
Object Assembly	11.11	2.66
Coding	10.61	2.53
WISC Verbal IQ	98.18	9.25
WISC Performance IQ	105.50	11.21
WISC Full Scale IQ	101.91	9.67

*See Appendix I for Botel Codes.

Table III

Mean Pre-Test Scores, By Method

Test Variable	Kinesthetic (N=31)	Auditory (N=38)	Visual (N=31)	Significant Effects (.05 Level)
WRAT--				
Reading G.E.	2.4	2.3	2.2	K ≥ V
Spelling G.E.	2.4	2.2	2.1	K > A & V
Gilmore--				
Accuracy G.E.	2.2	1.9	1.6	K > A & V
Comprehension G.E.	2.6	2.4	2.3	n.s.
Bender--Raw Score	3.2	2.5	2.4	n.s.
Benton--				
Error	8.9	9.9	8.6	n.s.
Correct	4.9	4.9	5.1	n.s.
Wepman--Raw Score	4.2	4.0	3.6	n.s.
Gilmore--Speed, WPM	72.2	60.6	59.5	K > A & V
MAT--				
Word Knowledge (2-68)*	2.2	2.2	2.2	n.s.
Word Discrimination (2-68)*	2.4	2.4	2.3	n.s.
Reading (2-68)*	2.1	2.1	2.0	n.s.
Word Knowledge (3-69)*	3.2	3.2	2.9	K & A > V
Reading (3-69)*	3.3	3.1	3.0	n.s.
Botel--				
Word Recognition	4.8**	3.4**	4.0**	K > A
Phonics Raw Score	48.7	38.7	39.2	K > A & V
Listening Raw Score	49.8	45.5	48.1	n.s.
Gray Oral--G.E.	2.2	1.9	1.8	K > A & V
WISC--				
Verbal IQ	100.2	96.5	98.1	n.s.
Performance IQ	104.4	103.9	108.5	n.s.
Full Scale IQ	102.6	100.1	103.5	n.s.
Lorge-Thorndike--IQ	97.6	100.8	103.2	V > K
Draw-A-Man--IQ	92.4	95.0	94.5	n.s.
PPVT--IQ	104.9	103.3	105.4	n.s.

*N=85

**See Appendix I for Botel Codes

Criterion Measure

SUMMARY

The post-test scores on reading tests administered at the end of the remediation year were specified as the criterion variable. Approaches to the criterion problem in other studies dealing with the impact of remediation are discussed.

The problem of criterion selection to assess change in reading level is an ongoing one. Despite their weaknesses (Krathwohl, 1966), difference scores are by far the most commonly used criteria. These difference scores are typically obtained on alternate forms of a group reading test, despite the problems encountered in interpreting group reading test scores because of the influence of the intelligence factor. Longitudinal studies of the effects of reading remediation used the following criteria: Balow (1965) does not name his criterion measures but discussed change in reading score; Lovell et al. (1963) used differences in Reading Age on the Schonell Graded Vocabulary Test and the Vernon Grade Word Reading Test; Bond and Fay (1950) compared grade equivalent gains made on "different forms of the Gates Tests," (the form was dictated by the pupil's reading level); Mouly and Grant (1956) used changes in grade equivalents on the California Reading Tests to devise a regression equation to predict to monthly gain in reading; Rasmussen and Dunne (1962) used change in reading scores between the California Achievement Test administered in sixth and the Iowa Silent Reading Test administered in ninth grade; Bliesmer (1962) compared "average yearly gains on the Reading subtest of the Metropolitan Achievement Tests and appropriate Gates Tests before and during reading remediation."

The criterion of "improvement" specified in this study was the reading level at the end of the remedial semester. The child's initial reading level was statistically controlled by including this variable among the predictor variables.

Reading level was reckoned by averaging the grade equivalent obtained by each child on the Wide Range Reading Test and the accuracy score of the Gilmore Oral Reading Test, both individually administered instruments. Averaging of the two test scores will probably increase the reliability of the measure obtained (Gulliksen, 1962). This is similar to the procedure typically utilized on group reading tests, e.g., the Gates Reading Survey yields scores in speed, vocabulary, and comprehension, with the total score being an average of the three grade equivalent scores.

Although word recognition is not the <u>only</u> variable of interest in measuring reading skills, most reading experts would agree that a child in the tird grade must emphasize word recognition skills. In order to comprehend a word, a child must first accurately perceive it, while speed is developed by practicing on materials which can be accurately read by a child. Given the average and above IQ's of the pupils involved, it would stand to reason that, as a group, their problems in learning to read consist of an inability to accurately recognize words. Nevertheless, the relationship of scores on the Gilmore and Gates Tests with the two word recognition scores (in context and in isolation) was inspected.

Chapter V

DERIVATION OF PREDICTION EQUATIONS AND SIMULATED ALLOCATION

SUMMARY

Prediction equations were calculated by means of stepwise regression analysis. For the approximately thirty subjects within each remedial method, the original variables were analyzed to determine which variables (and the weights of these variables) can best be combined to predict to change in reading scores. This yielded a series of three equations, one representing each remedial method. In this way, the pertinent variables for any child could be "plugged in" the three formulas for a quantitative prediction of improvement under each remedial method. This would allow for the future assignment of each child to the remedial method in which it is predicted he will benefit most.

Results of End of Year Testing, April - May, 1969

Individual testing with the Wide Range Achievement Test and the Gilmore Oral Reading Test was begun by the research analyst in the middle of April, 1969. The remedial teachers administered the Gray Oral Reading Test to their students during the last week of class in May, 1969. Attrition reduced the original sample from 102 to 100. All of these children were reached for final testing. Means and standard deviations for these tests are found in Table IV.

Table IV

Means and Standard Deviations of Test Variables

Test	Mean	S.D.
Final Wide Range Achievement Test, Reading, Grade Equivalent	3.19	.71
Final Gilmore Oral Reading Test, Accuracy, Grade Equivalent	3.36	.71
Final Gilmore Oral Reading Test, Comprehension, Grade Equivalent	3.38	.94
Final Gilmore Oral Reading Test, Speed, WPM	80.70	25.20

Test means broken down by remedial method are presented in Table V. The sample of 100 children includes 31 children in the kinesthetic method, 38 children in the auditory-phonic method, and 31 children in the visual method.

Table V

Means of Test Variables, By Method

Post-Test	Kinesthetic	Auditory	Visual
Wide Range Achievement Test, Reading	3.24	3.40	2.89
Gilmore Oral Reading Test, Accuracy	3.41	3.48	3.16
Gilmore Oral Reading Test, Comprehension	3.51	3.48	3.15

Utilizing the average WRAT Reading grade equivalent score and the Gilmore Accuracy grade equivalent score as the criterion measure, each child's improvement over the eight months of remedial reading was calculated. Table VI indicates that, overall, the auditory-phonic method was most successful for the greatest number of children. Analysis of variance indicated that the mean improvement was significantly greater for this method than for the other two methods.

Table VI

Mean Improvement Scores, By Method

$$\left(\text{May } \frac{\text{WRAT + Gilmore}}{2} \text{ Minus October } \frac{\text{WRAT + Gilmore}}{2}\right)$$

Method	Mean
Kinesthetic	1.00
Auditory-Phonic	1.32
Visual	1.12

This can be interpreted to mean that if the school district participating in this study were to select one remedial method as the method of choice for all its schools, the Auditory-Phonic method would be the most logical selection. However, it will be shown in this section that the use of three methods, allocated properly, is probably superior to using any one remedial method.

Development of Prediction Equations

The data for each remedial method required a series of separate analyses, since the purpose of each analysis was to discover the best single combination of independent variables to predict to the criterion for each remedial method, thus yielding a series of equations, each appropriate for predicting to change in reading score under one remedial treatment only. The equation for each method was based on data for children exposed to that method. Since approximately 100 children were involved, data for 31 to 38 children were used to derive each equation. This number should be sufficient to describe trends, this being consistent with the pilot study aspects of this project.

Prediction equations were established for the three remedial methods by means of a stepwise regression analysis* of the project variables. Only those independent variables were retained whose

---

*The mathematics of stepwise multiple regression as applied to the criterion variable -- final reading score -- can be presented briefly as follows:

Denoting y_j as the j^{th} child's reading score and $\hat{Y}_j = c_o + \Sigma_i \, c_i \, x_{ij}$ as a linear combination of the predictor variables (the x_i's) for that child, let $\hat{y}_j = a_o + \Sigma_i \, a_i \, x_{ij}$ represent the set of $\hat{Y}_j$'s for which $\Sigma_j \, (y_j - \hat{Y}_j)^2$ is a minimum. The set of $\hat{y}_j$'s then provides the best estimate in the sense of least squares of the observed y_j's.

With about 35 possible predictor variables, the equation $\hat{y}_j = a_o + \Sigma_j \, a_i \, x_{ij}$ would become quite cumbersome, and building that equation one variable at a time until the incremental predictive contribution of each added variable has an approximately one-in-ten or greater probability of resulting from chance effects would yield a less unwieldy equation and would, at the same time, eliminate redundancies in the information contained in the set of independent variables.

The first variable entered as a predictor is that variable most highly correlated with the criterion variable. Additional variables are then entered (and deleted) according to the extent to which they contribute to an increasing multiple regression coefficient, subject to a pre-specified significance level.

This very sketchy discussion of stepwise regression analysis can be supplemented by any of the more recent references in mathematical or applied statistics. A more thorough recapitulation of the theoretical considerations underlying stepwise regression analysis would only tend to add unnecessarily to the length of this section.

incremental contribution to prediction was significant at approximately the .10 level. Several runs were performed to assess the relative merits of part and whole scores (e.g., if the WISC Verbal Scale is a useful predictor, will the six subtest scores, each considered a separate predictor, be even more useful?).

The outcome of the analyses was, for each of the three techniques, a weighted sum of the characteristics predictive of success (as represented by a predicted score for the criterion measure). The weights assigned the characteristics differed from one equation to another since, as expected, the remedial techniques differed in their effects upon children. Also, some characteristics were specifically predictive for only one or two techniques. Thus the three equations contained not only different sets of predictive characteristics, but the weights accorded those characteristics common to one or more of the equations differed.

The prediction equations established in this study imply linear relationships between the predictive characteristics and the criterion variable. The possibility that curvilinear relationships may exist, and may be utilized to provide even better prediction, was examined. Departures from linearity over the range of the criterion variable were, however, found to be inconsequential. (This finding was further supported by those of a separate investigation by N. Silberberg and Iversen in 1968, in which quadratic and cubic models relating IQ to reading achievement represented no improvement over the corresponding linear model.)

Prediction Equations Developed on Final Testing (After Eight Months of Remedial Reading)

The Wide Range and Gilmore Reading Tests were administered to 100 children in the original sample and were utilized as the criterion variable.

The prediction equations yielded by this analysis (after eight months of remedial treatment) were:

KINESTHETIC — $R = .970$ — $R^2 = .942$ — S.D. = .1970

$\frac{\text{WRAT + Gilmore}}{2}$ =	.04796	(WISC, Arithmetic, Scaled Score)
	+ .05868	(WISC, Digit Span, Scaled Score)
	- .06438	(WISC, Picture Arrangement, Scaled Score)
	+ .03738	(WISC, Object Assembly, Scaled Score)
	+ .05800	(WISC, Coding, Scaled Score)
	+ 1.06120	(WRAT, Reading, Grade Equivalent)
	+ .11192	(Botel, Word Recognition, Coded Score)
	+ .04156	(C.A., 9/1/68, in months)
	- .00470	(D-A-M, IQ)
	- 4.81203	

AUDITORY-PHONIC $R = .875$ $R^2 = .766$ S.D. = .3580

$$\frac{\text{WRAT + Gilmore}}{2} = \begin{array}{rl} & .04172 \ \text{(WISC, Similarities, Scaled Score)} \\ + & .86190 \ \text{(WRAT, Reading, Grade Equivalent)} \\ - & .07912 \ \text{(Benton, Error, Raw Score)} \\ - & .26764 \ \text{(Benton, Correct, Raw Score)} \\ + & .48050 \ \text{(Gray Oral, Grade Equivalent)} \\ - & .02439 \ \text{(C.A., 9/1/68, in months)} \\ + & 4.60374 \end{array}$$

VISUAL $R = .902$ $R^2 = .814$ S.D. = .26007

$$\frac{\text{WRAT + Gilmore}}{2} = \begin{array}{rl} & .04154 \ \text{(WISC, Comprehension, Scaled Score)} \\ - & .11106 \ \text{(WISC, Object Assembly, Scaled Score)} \\ + & .36200 \ \text{(Gilmore, Accuracy, Grade Equivalent)} \\ + & .11122 \ \text{(Botel, Word Recognition, Coded Score)} \\ + & .01829 \ \text{(Botel, Listening Comprehension, Raw Score)} \\ + & .01077 \ \text{(Lorge-Thorndike, IQ)} \\ - & .32697 \ \text{(Sex; Male=1, Female=2)} \\ + & 1.38657 \end{array}$$

It will be noted that the prediction formulas yield correlation coefficients ranging from moderately to extremely high (.88 to .97) for each of the three methods. In addition, it should be noted that the test scores necessary for each of the three equations include group tests which are -- or could be -- routinely administered within the classroom to the children as well as some tests which are administered individually but do not require a great deal of testing time. Another interesting aspect of the prediction equations was the lack of emergence as powerful predictors of certain tests traditionally viewed as highly predictive of reading disability, including the Bender-Gestalt, the Benton, and the Wepman. This does not mean these tests were not correlated with the criterion variable, but rather that much of the information provided by these tests is also provided by those tests that emerged as predictors. In other words, the three equations based on this data would allow very workable and efficient predictions of a child's success in any one of the three remedial methods, utilizing variables which, aside from the WISC, would not be difficult to administer either in class or with little additional training.

The Use of Simulation Techniques

At this point, the simulation phase of the experiment was invoked. Utilizing the prediction formulas established on the groups to whom a specific remedial method was utilized, these prediction equations were then applied to the children in the other two groups.

The absolute difference between the predicted final reading score and the observed post-test was calculated. This was accomplished by including each child's tested characteristics to calculate what reading score he would have been predicted to get in the method to which he was actually assigned, and comparing this to his actual averaged score on the WRAT and Gilmore. The mean absolute differences, in which predicted scores are approximately only two months different on the average from actual scores, are shown in Table VII.

Table VII

Mean Absolute Difference Between
Predicted and Actual Reading Scores,
By Method and in Grade Equivalent Units

Method	Mean Absolute Difference
Kinesthetic	.136 G.E.
Auditory-Phonic	.218 G.E.
Visual	.203 G.E.

Realizing that the standard deviation (S.D.) becomes critical in assessing inter-group differences in a prediction problem such as this, the predicted scores were calculated to include values for a -.5 S.D. and a +.5 S.D. interval around the Predicted Value. In this way, inter-group differences can be more readily assessed. In our analyses, we arbitrarily defined one method as being predictively "best" for any given child if the grade equivalent for the value of -.5 S.D. in the superior method is higher than the grade equivalent for the value of +.5 S.D. for the inferior method. That is, the S.D. intervals must not overlap. In other words, using the first child on the list for the kinesthetic method (Table VIII-A) as an example, it can be seen that the values for the standard deviation intervals for all three methods overlap. Therefore, there is no one clearly defined "best" method for this child. It would appear, with some reasonable certainty, that he would succeed equally well if he were to be allocated to any one of the three methods.

However, looking at the second child on the list in Table VIII-A, it can be seen that the value for -.5 S.D. under the auditory-phonic method is at the 4.65 grade equivalent. This value is higher than the +.5 S.D. value under the other two methods (3.79 and 4.18, respectively) for that child. Therefore, if he were to be allocated to a remedial method, it would appear that the auditory-phonic method would be superior to the other methods <u>for him</u>. This approach (the use of standard deviation intervals) offers reasonable assurance that the allocation prodecure is less subject to chance errors than would occur if only comparisons were made of predicted scores under each method.

Table VIII-A. Children in Kinesthetic Method -- Predicted Scores Under Simulated Conditions for Each Child

Equation for Kinesthetic Method			Equation for Auditory-Phonic Method			Equation for Visual Method				
-.5SD	PRED. VALUE	+.5SD	-.5SD	PRED. VALUE	+.5SD	-.5SD	PRED. VALUE	+.5SD	October, 1968 WRAT-GIL/2	May, 1969 WRAT-GIL/2
3.28	3.38	3.48	3.42	3.60	3.78	3.19	3.32	3.45	2.1	3.2
3.60	3.69	3.79	4.65	4.83	5.01	3.92	4.05	4.18	2.7	4.0
3.12	3.21	3.31	3.29	3.46	3.64	3.90	4.03	4.16	2.5	3.3
3.04	3.14	3.24	3.11	3.29	3.47	3.11	3.24	3.37	2.1	3.1
3.11	3.21	3.31	2.69	2.87	3.04	3.34	3.47	3.60	2.0	3.2
1.49	1.59	1.69	2.47	2.65	2.83	3.10	3.23	3.36	1.2	1.3
2.12	2.22	2.32	2.79	2.97	3.15	2.36	2.49	2.62	1.4	2.2
4.23	4.33	4.43	3.93	4.11	4.29	3.27	3.40	3.53	3.0	4.3
1.99	2.08	2.18	2.91	3.08	3.26	2.91	3.04	3.17	2.0	2.4
2.62	2.72	2.82	2.06	2.24	2.42	2.91	3.04	3.17	1.8	2.8
2.78	2.88	2.98	3.57	3.75	3.92	3.36	3.49	3.62	2.4	2.7
3.82	3.92	4.02	3.58	3.76	3.94	4.20	4.33	4.46	2.7	3.6
3.34	3.44	3.54	3.21	3.39	3.57	2.88	3.01	3.14	2.4	3.1
3.61	3.70	3.80	4.05	4.23	4.41	3.88	4.01	4.14	2.6	4.0
3.56	3.66	3.75	3.60	3.78	3.96	5.11	5.24	5.37	2.8	3.6
3.41	3.51	3.61	4.00	4.18	4.36	3.08	3.21	3.34	1.9	3.6
3.77	3.87	3.97	3.66	3.84	4.02	4.09	4.22	4.35	2.4	3.8
3.38	3.48	3.58	3.98	4.16	4.34	4.30	4.43	4.56	2.6	3.4
3.31	3.40	3.50	4.51	4.69	4.87	3.59	3.72	3.85	2.7	3.3
3.86	3.96	4.05	4.66	4.84	5.02	3.36	3.49	3.62	2.9	3.8
2.55	2.65	2.75	2.95	3.13	3.31	3.10	3.23	3.36	1.6	2.7
2.91	3.01	3.11	3.12	3.30	3.48	2.81	2.94	3.07	2.1	3.1
3.32	3.42	3.52	4.62	4.80	4.98	3.00	3.13	3.26	2.1	3.2
4.22	4.32	4.42	4.25	4.43	4.61	3.31	3.44	3.57	3.0	4.2
4.01	4.11	4.21	4.18	4.36	4.54	3.61	3.74	3.87	2.8	4.2
4.08	4.18	4.28	4.45	4.63	4.81	3.86	3.99	4.12	3.2	4.3
3.26	3.36	3.46	3.32	3.49	3.67	2.79	2.92	3.05	2.1	3.6
3.07	3.17	3.26	3.88	4.06	4.24	2.89	3.02	3.15	2.2	3.1
3.86	3.96	4.06	3.88	4.06	4.24	2.60	2.73	2.86	2.2	3.9
3.08	3.18	3.28	3.33	3.51	3.68	3.04	3.17	3.30	2.4	3.1
2.28	2.38	2.48	3.62	3.80	3.97	2.45	2.58	2.71	1.6	2.4

Table VIII-B. Children in Auditory-Phonic Method -- Predicted Scores Under Simulated Conditions for Each Child

Equation for Kinesthetic Method			Equation for Auditory-Phonic Method			Equation for Visual Method				
-.5SD	PRED. VALUE	+.5SD	-.5SD	PRED. VALUE	+.5SD	-.5SD	PRED. VALUE	+.5SD	October, 1968 WRAT-GIL/2	May, 1969 WRAT-GIL/2
2.40	2.50	2.60	3.37	3.55	3.73	3.21	3.34	3.47	2.2	3.8
2.70	2.80	2.90	3.23	3.41	3.59	3.39	3.52	3.65	2.0	3.4
3.57	3.67	3.77	3.62	3.80	3.98	2.68	2.81	2.94	2.4	3.9
3.76	3.85	3.95	4.35	4.53	4.70	3.23	3.36	3.49	2.1	4.5
3.51	3.61	3.71	3.08	3.25	3.43	3.39	3.52	3.65	1.8	3.4
3.86	3.96	4.06	2.22	2.40	2.58	3.00	3.13	3.26	1.5	2.4
3.10	3.20	3.30	3.73	3.91	4.09	3.76	3.89	4.02	2.5	4.2
99	1.09	1.18	2.17	2.34	2.52	2.68	2.81	2.94	1.2	1.7
2.43	2.53	2.63	3.12	3.30	3.48	2.56	2.69	2.82	2.2	3.4
3.12	3.22	3.32	3.64	3.82	4.00	3.41	3.54	3.67	2.1	3.7
3.79	3.89	3.99	4.25	4.43	4.61	4.00	4.13	4.26	2.5	4.1
2.99	3.09	3.19	2.96	3.14	3.31	3.73	3.86	3.99	1.8	3.1
1.76	1.86	1.95	2.65	2.83	3.01	2.59	2.72	2.85	1.7	3.2
3.40	3.49	3.59	3.20	3.38	3.56	3.03	3.16	3.29	2.3	3.6
3.11	3.21	3.31	3.87	4.05	4.22	3.09	3.22	3.35	2.8	4.0
3.62	3.72	3.82	3.41	3.59	3.76	2.85	2.98	3.11	2.6	3.3
1.66	1.76	1.86	3.09	3.27	3.45	3.10	3.23	3.36	1.9	3.1
3.03	3.13	3.23	3.37	3.55	3.73	2.97	3.10	3.23	2.3	3.8
1.99	2.09	2.19	3.19	3.37	3.54	2.23	2.36	2.49	1.5	2.9
2.72	2.82	2.91	3.05	3.23	3.41	2.33	2.46	2.59	1.5	3.2
2.62	2.72	2.82	3.25	3.43	3.60	2.49	2.62	2.75	2.5	3.5
3.39	3.49	3.59	3.81	3.99	4.17	3.40	3.53	3.66	2.9	4.1
2.65	2.74	2.84	3.03	3.21	3.39	2.70	2.83	2.96	2.0	3.3
3.28	3.38	3.48	4.12	4.30	4.48	3.16	3.29	3.42	2.8	4.3
2.80	2.90	3.00	2.99	3.17	3.34	2.96	3.09	3.22	2.1	4.5
1.93	2.02	2.12	2.72	2.90	3.08	3.17	3.30	3.43	2.1	2.6
4.37	4.47	4.57	4.10	4.27	4.45	3.37	3.50	3.63	2.6	4.2
4.50	4.60	4.69	3.92	4.10	4.28	3.35	3.48	3.61	2.4	3.9
2.71	2.81	2.90	2.93	3.11	3.29	2.19	2.32	2.45	1.5	2.9
2.38	2.48	2.58	3.22	3.40	3.58	2.47	2.60	2.73	1.5	3.6
2.00	2.10	2.20	2.51	2.69	2.87	2.06	2.19	2.32	1.4	2.5
2.72	2.81	2.91	2.71	2.89	3.07	2.92	3.05	3.18	2.3	2.3
2.38	2.47	2.57	3.33	3.51	3.69	3.36	3.49	3.62	2.1	3.3
2.72	2.82	2.92	2.20	2.38	2.56	1.68	1.81	1.94	1.4	2.5
3.49	3.59	3.68	3.58	3.76	3.94	2.72	2.85	2.98	2.6	3.8
2.43	2.52	2.62	1.99	2.17	2.35	1.77	1.90	2.03	1.3	2.2
5.28	5.38	5.48	4.09	4.27	4.45	3.66	3.79	3.92	3.0	3.9

Table VIII-C. Children in Visual Method -- Predicted Scores Under Simulated Conditions for Each Child

Equation for Kinesthetic Method			Equation for Auditory-Phonic Method			Equation for Visual Method				
-.5SD	PRED. VALUE	+.5SD	-.5SD	PRED. VALUE	+.5SD	-.5SD	PRED. VALUE	+.5SD	October, 1968 WRAT-GIL/2	May, 1969 WRAT-GIL/2
2.57	2.67	2.77	3.38	3.55	3.73	2.97	3.10	3.23	1.6	3.4
3.53	3.63	3.73	2.18	2.36	2.54	2.58	2.71	2.84	1.5	2.6
4.07	4.17	4.27	3.63	3.81	3.99	3.88	4.01	4.14	2.7	4.1
2.61	2.71	2.81	2.79	2.97	3.15	2.56	2.69	2.82	1.7	3.0
3.35	3.45	3.55	3.63	3.81	3.99	2.81	2.94	3.07	1.5	2.8
3.37	3.47	3.57	3.18	3.36	3.54	3.04	3.17	3.30	2.5	2.8
2.64	2.74	2.84	3.12	3.30	3.48	2.51	2.64	2.77	1.5	2.8
1.52	1.62	1.72	2.08	2.26	2.44	2.26	2.39	2.52	1.3	2.1
2.70	2.80	2.90	2.73	2.91	3.09	3.24	3.37	3.50	1.8	3.1
3.87	3.96	4.06	3.55	3.73	3.91	3.22	3.35	3.48	2.6	3.2
2.63	2.73	2.83	1.37	1.54	1.72	2.05	2.18	2.31	1.2	1.9
2.27	2.37	2.47	3.67	3.85	4.02	3.31	3.44	3.57	2.0	3.4
3.33	3.43	3.53	3.42	3.60	3.78	3.40	3.53	3.66	2.5	3.8
2.50	2.60	2.70	2.96	3.14	3.31	2.63	2.76	2.89	1.5	2.7
2.94	3.04	3.14	3.71	3.89	4.07	2.93	3.06	3.19	1.9	3.4
2.25	2.34	2.44	3.07	3.24	3.42	2.51	2.64	2.77	1.5	2.8
2.18	2.28	2.38	2.57	2.75	2.93	2.48	2.61	2.74	1.4	2.3
3.51	3.60	3.70	3.39	3.57	3.75	3.68	3.81	3.94	2.6	3.5
3.22	3.31	3.41	3.09	3.27	3.45	2.62	2.75	2.88	1.9	2.6
3.11	3.21	3.30	2.85	3.03	3.21	3.20	3.33	3.46	1.6	3.2
3.19	3.28	3.38	3.20	3.37	3.55	3.06	3.19	3.32	1.9	3.3
3.07	3.17	3.27	2.90	3.08	3.26	2.81	2.94	3.07	2.5	3.1
2.84	2.94	3.04	2.89	3.06	3.24	2.14	2.27	2.40	1.5	2.1
3.55	3.65	3.74	4.44	4.62	4.80	3.70	3.83	3.96	2.2	3.4
3.69	3.79	3.89	3.69	3.87	4.05	3.01	3.14	3.27	2.1	3.2
3.19	3.29	3.39	3.19	3.37	3.55	2.82	2.95	3.08	1.8	3.0
3.14	3.24	3.34	3.12	3.30	3.48	3.21	3.34	3.47	2.1	3.5
2.17	2.27	2.37	2.60	2.78	2.96	2.41	2.54	2.67	1.9	2.4
3.43	3.53	3.63	4.21	4.39	4.56	3.34	3.47	3.60	2.4	3.4
2.37	2.47	2.57	2.96	3.14	3.32	2.14	2.27	2.40	1.5	2.6

Another arbitrary factor which went into our classification scheme was to identify those children for whom a remedial method was "wrong." A child was classified as not responding to a remedial method if the predicted value under a remedial method was less than eight months higher than his pre-test score. Since the children had been exposed to their remediation for eight months, we established as our criterion for "success" in a method that the child progressed at least at a normal rate for these eight months. Possibly, this requirement was somewhat stringent in that all of the children involved in the research project had not been able to progress in reading at a normal rate before the initiation of this project. Nevertheless, we felt that this was a necessary additional classification. Also added in the "wrong" category were those methods remaining when another method was clearly deemed "right" for a child. (If visual is the single "best" method for a child, phonic and kinesthetic are "wrong" methods for him.)

This consideration of normal progress in learning rate required another classification in our allocation scheme. In some cases, the range of values around a child's predicted scores (±.5 S.D.) on two or three methods overlapped and thus could be viewed as equal. However, if one or two of the overlapping methods yielded a mean predicted score of less than eight months higher than that child's actual pre-test score, some consideration of this fact had to be made. In the cases of two or three equally "best" methods, but where one or two of these methods yielded a mean improvement of less than .8, the method or methods with a mean improvement score of .8 or better was classified as "better," even though it overlapped with an unsuccessful method. Of course, the remaining method was included in the category of "wrong" method, even though it yielded predicted values significantly higher than another "wrong" method.

To summarize: If the mean predicted improvement for a method was less than .8, it was considered "wrong." If the mean predicted improvement was .8 or better, that method was considered "best," (if two or three "best" methods met this criterion and overlapped, there could be two or three "best" methods for one child). If the value for the +.5 S.D. for a method overlapped with a method that yielded a mean predicted increase of less than .8, this successful method (or methods) was considered "better," in order to distinguish it from a "best" method, which was more clearly successful for that child.

For some children, no remedial method yielded predicted increases of .8 or more. Two children fell into this classification. In other words, there are two children who would not have predictively responded in an efficient manner to any of the three methods, although 17 children improved less than eight months in the method to which they were randomly allocated. Possibly this research project will, for the first time, identify those children for whom a more radical educational innovation is required (e.g., N. Silberberg and M. Silberberg, 1969b).

In assigning children to "right" methods for them, three possibilities existed: either one method, two methods, or all three methods were predicted "best" or "better" for him. In all cases, this involved the predictive score for the "right" method or methods being eight or more months higher than the pre-test value. For some children, such predictions occurred in only one method, while for other children such predictions occurred for two and three methods. Table IX presents the breakdown on the potential allocation options available based on the predictive data assembled in the analysis.

Table IX

Allocation of Children by Prediction Equations

Number of Children With 0 "Best" or "Better" Methods	2
Number of Children With 1 "Best" or "Better" Method	56
Number of Children With 2 "Best" or "Better" Methods	30
Number of Children With All 3 Methods Equally Efficient	12

Table X shows the effects of simulated prediction and the outcomes which "could have been" had proper allocation taken place. The "right" method included both "best" and "better" methods for each child.

To illustrate, let us examine the auditory method results. The 38 children who were in classes where this method was taught gained 1.32 years in eight months between pre-test and follow-up. However, 28 children were properly (according to the prediction equations) placed, by chance, in this method and they actually gained 1.47 years, on the average. Had the prediction equations been available before beginning remediation, it would have been predicted that these 28 children would have gained 1.41 years, on the average, in this period. The other 10 children, for whom this method was not demed the "right" method, gained only an average of .90 years in eight months. According to the prediction equations, however, 9 of this wrongly placed group of children would have gained 1.88 years if placed in another method deemed "right" for each child. (For one child in this original group of 38, none of the methods would have appeared to help.) Almost a full year of additional gain might have been realized by those children "incorrectly" allocated to one of the other two methods had they been properly allocated to a "right" method. Similarly, an average of seven to eight months might have been gained in the other two methods by properly re-allocating 23 and 19 children, respectively.

Table X

Effects of Simulative Prediction

Mean Gain	Kinesthetic	Auditory	Visual
ACTUAL Mean Gain, all S's in actual classes	1.00 (N=31)	1.32 (N=38)	1.12 (N=31)
ACTUAL Mean Gain, S's in "right" method	1.24 (N=8)	1.47 (N=28)	1.30 (N=11)
PREDICTED Mean Gain, S's in "right" method	1.29 (N=8)	1.41 (N=28)	1.28 (N=11)
ACTUAL Mean Gain, S's in "wrong" method	.91 (N=23)	.90 (N=10)	1.02 (N=20)
PREDICTED Mean Gain, if S's in "wrong" method had been placed in "right" method	1.70 (N=23)	1.88 (N=9)	1.68 (N=19)

It would appear that the total gain from the proper allocation of children, utilizing simulation techniques, would have been very great. Children who were, by chance, placed in the "wrong" remedial reading method would have achieved anywhere from one-half year to one year higher in reading, had they been placed properly. Such decisions could have affected 51 percent of the children in the sample. In other words, it would certainly appear that the procedures outlined in this experiment could be expanded and utilized in the allocation of children to remedial reading methods. Of course, larger samples are required and additional remedial reading techniques would also be useful. Quite possibly, when data is available on a sufficient number of children and a sufficient number of methods, the effects of remedial reading could be more positive than have occurred (N. Silberberg and M. Silberberg, 1969a).

Chapter VI

SUMMARY AND CONCLUSIONS

Typically, attempts at individualizing remedial instruction in the past have either allowed different children to work with the same instructional materials at a different pace, or involved an a priori diagnosis of a child's difficulties and a resultant remedial program based on postulated logical relationships between the diagnostic characteristics and the remedial materials. Longitudinal studies (N. Silberberg and M. Silberberg, 1969a) have not demonstrated a permanent effect of the gains made in these remedial programs.

This study was a first attempt to utilize simulation techniques in assigning remedial reading to children. One hundred children in one suburban St. Paul, Minnesota, school district were involved. These children represented about 10 percent of the total third grade population in that district. All subjects were included in a remedial reading program based on their having scored 90 or above on a group IQ test administered in second grade, and being designated as reading one grade below placement by their second grade classroom teachers.

The children were placed in sixteen remedial classes in twelve different schools, under the supervision of seven experienced remedial reading teachers. The classes were divided into three groups, each group providing a different remedial method to its members. An attempt was made to equalize the teacher effect across groups.

The three remedial methods were selected to be as distinct, one from another, as possible. Each emphasized a different sensory input, and included a visual, a kinesthetic, and an auditory-phonic approach. Each child in the sample received eight months remedial instruction under the one remedial method assigned to his class.

Prior to the initiation of the remedial instruction, each child was tested with a battery of individually administered tests, including tests purportedly tapping skills in the perceptual, perceptual-motor, and verbal areas. In addition, each child was individually tested on three reading tests. Also included in the data pool was information on previously administered group IQ and reading tests, as well as quantified demographic characteristics. These approximately forty scores, together with WISC scaled scores obtained by testing throughout the year, comprised the pre-test data pool.

Upon completion of the eight months of remedial instruction, each child was again administered two reading tests. The average of these two tests was utilized as the criterion measure.

Stepwise regression techniques were then utilized to predict from the scores in the pre-test data pool to the criterion measure for each child. Multiple correlation coefficients ranging from .875 to .970 were obtained in these equations. The difference between each child's actual score on the criterion measure and that predicted from his pre-test data differed, on the average, by less than two months.

Assuming that a child with a set of tested characteristics in one class is no different than another child with the same characteristics in another class, simulation procedures were then invoked. This involved applying a child's test scores to the prediction equations for the two methods other than the one his class used. In this way, each child's reaction to the other two methods could be simulated and compared to his predicted and actual scores on the criterion measure in his own method.

It was found that more than half the total sample predictively would have achieved a tested reading level of from seven months to one year higher had they been placed in methods other than the one to which, by chance, they had been assigned.

Thus, the purpose of this pilot project has been achieved. It would appear that this study has demonstrated that children's reactions to a remedial treatment can be predicted on the basis of their individual characteristics. It has also provided a model for developing research which utilizes simulation techniques in individualizing instruction. The questions of the sample-to-sample reliability of the prediction equations, whether including more children and a greater choice in remedial methods would make prediction more precise, and whether the effects of remediation for those children who were properly placed by chance in their "right" method do not dissipate within one or two years of terminating remedial instruction -- all must await replication and follow-up of this study.

REFERENCES

Balow, B. The long term effect of remedial reading instruction. Reading Teacher, 1965, 18, 581-586.

Bliesmer, E. Evaluating progress in remedial reading programs. Reading Teacher, 1962, 15, 344-350.

Bliesmer, E. and Yarborough, B. A comparison of 10 different beginning reading programs in first grade. Phi Delta Kappan, 1965, 46, 500-504.

Bond, G. and Fay, L. A report of the University of Minnesota Reading Clinic, Journal of Educational Research, 1950, 43, 385-390.

Fitzgerald, L. Some effects of reading ability on group intelligence test scores in the intermediate grades. Unpublished doctoral dissertation, University of Iowa, 1960.

Gulliksen, H. Theory of Mental Tests. New York: John Wiley, 1962.

Haring, N. and Ridgeway, R. Early identification of children with learning disabilities. Exceptional Children, 1967, 33, 387-398.

Heckelman, R. A neurological impress method of reading instruction. Mimeographed report, Merced County Schools: Merced, California, 1962.

Krathwohl, D. How to Prepare a Research Proposal. Syracuse: Syracuse University Bookstore, 1966.

Lovell, K., Byrne, C., and Richardson, B. A further study of the educational progress of children who received remedial education. British Journal of Educational Psychology, 1963, 33, 3-7.

Lytton, H. Follow-up of an experiment in selection for remedial education. British Journal of Educational Psychology, 1967, 37, 1-9.

Mouly, G. and Grant, V. A study of the growth to be expected of retarded readers. Journal of Educational Research, 1956, 49, 461-465.

Rasmussen, G. and Dunne, H. A longitudinal evaluation of a junior high school corrective reading program. Reading Teacher, 1962, 34, 363-396.

Silberberg, N. and Feldt, L. An investigation of the intellectual and perceptual correlates of reading difficulties. Journal of School Psychology, Summer, 1968, 6(4), 237-245.

Silberberg, N. and Iversen, I. The effects of kindergarten instruction in letters and numbers on end of first grade reading, Final Report. U. S. Office of Economic Opportunity, 1968, 65 pages.

Silberberg, N., Iversen, I., and Silberberg, M. A model for classifying children according to their reading level. Mimeographed report, Kenny Rehabilitation Institute: Minneapolis, 1969.

Silberberg, N. and Silberberg, M. Myths in remedial education. Journal of Learning Disabilities, April, 1969, 2(4), 209-217. (a)

Silberberg, N. and Silberberg, M. The bookless curriculum: An educational alternative. Journal of Learning Disabilities, June, 1969, 2(6), 302-307. (b)

APPENDIX I

SAMPLE CALCULATION OF SIMULATION

SAMPLE CALCULATION OF SIMULATION

Simulating Child No. 101, who was assigned to the Auditory Method, his response to the Kinesthetic Method would be calculated thusly:

$$
\begin{aligned}
\hat{y} = \; & .04796 \times 13 = + .62348 \\
+ \; & .05868 \times 7 = + .41076 \\
- \; & .06438 \times 11 = - .70818 \\
+ \; & .03738 \times 10 = + .37380 \\
+ \; & .05800 \times 6 = + .34800 \\
+ \; & 1.06120 \times 2.2 = + 2.33464 \\
+ \; & .11192 \times 3 = + .33576 \\
+ \; & .04156 \times 97 = + 4.03132 \\
- \; & .00470 \times 94 = - .44180 \\
& \qquad - 4.81203 \\
& \qquad \overline{2.49576} = 2.50
\end{aligned}
$$

Using the same child's scores in the equation predicting his response to the Auditory Method:

$$
\begin{aligned}
\hat{y} = \; & .04172 \times 13 = + .54236 \\
+ \; & .86190 \times 2.2 = + 1.89618 \\
- \; & .07912 \times 15 = - 1.18680 \\
- \; & .26764 \times 3 = - .80292 \\
+ \; & .48050 \times 1.8 = + .86490 \\
- \; & .02439 \times 97 = - 2.36583 \\
& \qquad + 4.60374 \\
& \qquad \overline{3.55163} = 3.55
\end{aligned}
$$

Using the same child's scores in the equation predicting his response to the Visual Method:

$$
\begin{aligned}
\hat{y} = \; & .04154 \times 6 = + .24924 \\
- \; & .11106 \times 10 = - 1.11060 \\
+ \; & .36200 \times 2.3 = + .83260 \\
+ \; & .11122 \times 3 = + .33366 \\
+ \; & .01829 \times 49 = + .89621 \\
+ \; & .01077 \times 100 = + 1.07700 \\
- \; & .32697 \times 1 = - .32697 \\
& \qquad + 1.38657 \\
& \qquad \overline{3.33771} = 3.34
\end{aligned}
$$

These results are on the first line of Table VIII-B, page 25. Noting the standard deviation range for this child, he would be allocated into either the Auditory or Visual Method, each of these being deemed as "right" for him.

The Botel Word Recognition Test is entered into the equation based on the following codes:

P.P. = 01
P. = 02
1 = 03
1^1 = 04
1^2 = 05
2^1 = 06
2^2 = 07
3^1 = 08