1969

An evaluation of three methods of instruction for teaching sight vocabulary to educable mentally handicapped children

Betty Jean (Wolter) Wilson
Southern Illinois University Edwardsville

AN EVALUATION OF THREE METHODS OF INSTRUCTION

FOR TEACHING SIGHT VOCABULARY TO

EDUCABLE MENTALLY HANDICAPPED CHILDREN .

by (Wolter)

Betty Jean Wilson

Bachelor of Science Degree
in Psychology, 1966
Southern Illinois University
Edwardsville, Illinois

A Thesis Submitted in Partial

Fulfillment of the Requirements for

the Master of Science Degree

Department of Psychology
in the Graduate School
Southern Illinois University

June 1969 .

SOUTHERN ILLINOIS UNIVERSITY

The Graduate School

May 5 , 19 69

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION

BY Betty Jean Wilson

ENTITLED Three Methods of Teaching Sight Vocabulary to Educable Mentally Handicapped Children

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF Master of Science

Eva D Ferguson
Thesis Director

Erwin H Brinkmann
Faculty Chairman

Edw. 3/66

ACKNOWLEDGEMENTS

Acknowledgement is extended to Dr. Ewin H. Brinkmann and Dr. Orval G. Johnson for their valuable suggestions concerning the planning of the experiment and the preparation of the report. A special appreciation is also given to Dr. Eva D. Ferguson, primary advisor, who provided the patient counsel, encouragement and support necessary for completion of this thesis.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Definition and Description of the Problem

Investigations of educational problems of the mentally retarded indicate that, in general, retarded children read[1] below their mental ages. This condition is especially true of children placed in special public school classes for the mentally retarded. Since the ability to read has great practical and educational significance in the progress of the mentally retarded, research directed toward the discovery of more effective teaching methods is very important. It was the purpose of this study, therefore, to compare three methods of teaching sight vocabulary. With each method the children worked without the direct aid of a teacher. However, one method may be described as traditional because the task required the use of paper and pencil, and the student received no immediate knowledge of results. The second and third methods featured a teaching device which gave the student automatic feedback.

Research on Special Classes for the Mentally Retarded

The number of special education classes for mentally retarded children has risen steadily since the organization of the first classes in public schools in 1896 (Kirk, 1964, p. 57). These classes have been established on the logical belief that smaller classes and tasks related to ability level rather than age would enhance academic

---

[1]Russell and Fea (1963, p. 868) define reading as (1) identifying the symbol, and (2) obtaining meaning from the recognized symbol. The experiment described in this paper was an investigation of only one of the processes involved in reading, namely, "identifying the symbol."

performance. However, with one recent exception, numerous investigations have shown that academic achievement of educable mentally retarded children in special classes is either not significantly different or is inferior to that of retarded children remaining in the regular grades (Bennett, 1932; Blatt, 1958; Cassidy & Stanton, 1959; Wrightstone, Forlano, Lepkowski, Sontag, & Edelstein, 1959; Mullen & Itken, 1961; Warren, 1962). A survey of this research indicates methodological inadequacies including (a) lack of control over sample selection, (b) effects of varying school experiences prior to special class placement, and (c) lack of specification and control over the special class educational program.

The only study (Goldstein, Moss, & Jordan, 1964) in which superiority of special classes was demonstrated has two important facets. First, it was designed and executed to answer criticism of previous studies. As stated by the authors, "The methodology used in this study should obviate the necessity for further studies... (p. iv)." The methodology referred to included testing of approximately 2000 children who were ready to enter the first grade, random assignment of subjects to regular and special classes, and supervision of the special class educational program by a member of the investigating team. In addition, the teachers had degrees in the education of the mentally retarded. The second important facet of the study is that, in compensating for the inadequacies of other studies, the investigators created special classes which were not representative of special education classes established in public schools. Usually educable retarded children in special classes have an age range of from 3 to 5 years, while in the Goldstein study the subjects were the same age.

Also, eliminating previous school experience as a variable caused a major difference because, after one year of schooling, many of the subjects in both groups increased in IQ sufficiently to declassify them as mentally retarded even with a criterion of 85 IQ and below. Upon completion of the four year study only 19 of the 41 children remaining in the special classes and 15 of the 52 children in regular grades had IQ's of 80 or below. Since no significant differences were found between the groups as a whole, separate analyses were made for children with terminal IQ's of 81 or above, and for those with IQ's of 80 or below. The differences between the high IQ groups were not significant. Among the low IQ groups, however, the experimental group tended to exceed the control group in arithmetic computation and problem solving, and in basic social knowledge. The results of the experiment with regard to reading and language achievement were equivocal with three of ten subtests showing a significant difference in favor of the special classes. The authors conclude that the study "...indicate[s] that the special class is beneficial to children whose IQ's are 80 and below (p. 7)." It is doubtful, however, that the results of this study can be generalized to other special classes. The investigation does, though, offer promise that academic performance of educable mentally handicapped children can be increased with better teaching methods.

Description of Special Classes

An examination of children in special classes reveals a relatively large range of age and performance levels. Since special classes are established primarily to give more attention to students on a level commensurate with their ability, small groups within a class meet with the teacher for reading and arithmetic instruction. Classes may have as many as five reading groups and two or three for arithmetic. A child may spend as much as an hour a day in separate groups. Consequently, however, he must also work individually for a much greater length of time while the teacher is engaged with other groups. Because the advancement of a mentally retarded child is comparatively slow, the assignments are, of necessity, repetetive. The teacher marks the papers and returns them; perhaps she has him correct his mistakes, perhaps not. Possibly the child does not look at the papers after they have been corrected. This procedure gives the students many opportunities to compound errors. That making wrong responses before making correct ones has deletorious effect on later performance is indicated in the work of Kaess & Zeaman (1960), Knight (1963-1964), and Peterson & Brewer (1963) with normals. Perhaps one of the factors contributing to the higher achievement of the retarded in regular classes is related to the greater amount of time spent in total class instruction. In this situation the child has an opportunity to compare his knowledge with the correct response, and therefore, a measure of immediate feedback is available.

Research with Methods Giving Immediate Feedback

Programmed texts giving immediate feedback and/or reinforcement to the pupil working individually would tend to eliminate the condition just described. However, there are no published accounts of research with retardates using programmed texts. A few studies are available which compare conventional classroom teaching of reading with methods that give immediate feedback (such as teaching machines), but the results are inconclusive. Of these studies, the most carefully controlled was conducted by Blackman & Capobianco (1965), in which the same individuals served as both teachers and experimenters. Performance of teaching machine and classroom Ss improved to a statistically significant extent over the one year evaluation period. There were, however, no differences between the two groups. A retention test administered 3½ months after training showed no loss for either group.

Malpass, Gilmore, Hardy, and Williams (1964) compared four approaches to teaching reading and spelling: a multiple choice device, a typewriter keyboard, a conventional classroom, and a tutorial procedure. The classroom procedures were uncontrolled. Teachers were simply given a list of the programmed words and encouraged to present them during their regular classroom sessions. The gains made by the tutorial group matched those made by the machine groups, and all t-ratios between the latter experimental conditions and the classroom condition significantly favored the experimental groups.

Another investigation (Vergason, 1966), compared traditional classroom presentation of words to automated slide projector instruction. The automated instruction was similar to the paired-anticipation method used in verbal learning experiments. The classroom teachers taught the words in a variety of ways. Both methods produced good retention rates with Ss not differing on retention after one day but showing significant differences at one, two, and fourteen months in favor of auto-instruction. Vergason concludes that the traditional method may "...produce better understanding...", but "it may also introduce elements of distraction or confusion." On the other hand, "...the automated method might be more effective because it is highly consistent (p. 688)."

A study of reading instruction under either a programmed instruction condition, a standard classroom condition (not described), a procedure in which programmed instruction and classroom sessions were alternated, and a control condition of no instruction was conducted by Ellson, Engle, Barber, and Kempworth (1962). The programmed instruction condition required the presence of a tutor to reinforce and correct the student's oral responses. Inter-group comparisons showed both the programmed instruction and alternation groups to have acquired significantly more words than the control group, and the alternation Ss to have acquired significantly more words than the classroom Ss.

At least two authors (Eigen, 1965; Stolurow, 1962) have argued against research comparing programmed instruction with conventional teaching. The major criticism is that neither variable can be clearly defined; there is no one method of programmed instruction, just as there is no conventional teaching method. An additional argument which could be advanced is that "classroom instruction," when used as a variable, is difficult to control. Unless the same individuals serve as teachers and experimenters there is no assurance that the teachers will present the material that is programmed.

The arguments against a comparison of programmed instruction and conventional teaching methods are cogent. However, in the practical realm and especially for mentally retarded classes, the greater efficiency of one method over the other could be extremely important.

Therefore, attempts to compare the two methods should not be dropped. Studies can be designed which compensate for the methodological difficulties of the problem. One way of dealing with the definition and control of "conventional classroom methods" is to compare only one of the many varieties of teaching aids with an instrument which gives immediate feedback. It may be recalled, for instance, that mentally retarded children in special classes spend a disproportionate amount of time doing "seatwork." The types of individual reading assignments, especially on the lower grade levels, are standard, being mainly derived from published workbooks. The tasks consist of filling in the correct word or letter(s), making a mark on the correct picture, underlining the correct word, drawing a line from a picture to the identifying word, etc. The relative effectiveness of any or all of these tasks in which the student works without the direct aid of a teacher can be compared to an instrument (or text) which gives automatic feedback. In an experiment of this kind the classroom method can be defined and controlled without the variables of teacher personality and techniques interfering. Also, since only one feature of programmed instruction is incorporated in the experiment (i.e., immediate feedback), sequencing and "program" would not be confounding variables.

Paired Programmed Learning

One assumption that appears to underlie the use of the typical automated or semi-automated teaching instrument is that the student is interested and, therefore, attending to the task. However, with some of the teaching machines it is possible for the subject to complete the program without ever having attended to the subject matter. A random pressing of buttons until the correct button is touched will often advance the machine to another frame. Thus it is possible for the subject's attention to be focused on the buttons rather than on the screen. One of the arguments for paired programmed learning with normals is based on the factor of greater interest resulting from interaction. Other arguments which have been advanced include: (a) it is more economical, (b) students help each other over difficulties, and (c) retention is better for paired students than for individuals. Two studies with normals (Dick, 1963; Amaria, 1967) offer evidence supporting greater retention by pairs, but the results are confounded by variables such as personality and ability of the paired subjects. Other studies indicate that learning in pairs is not detrimental (Austwick, 1965; Sawaris, 1966; Amaria, Biran, & Leath, 1966). These investigations suggest that paired learning of retardates might be a satisfactory method of encouraging attention to the task.

Empirical Hypotheses

Two hypotheses growing out of the previous discussion were tested in the present experiment.

1. Retardates receiving immediate feedback from a teaching instrument achieve a greater reading vocabulary than Ss working individually on paper and pencil tasks that do not give immediate knowledge of results.

2. Paired retardates working with a teaching instrument that gives immediate feedback achieve a greater reading vocabulary than individual *S*s working with the same instrument.

CHAPTER II

Method

Training Task

The training task consisted of learning to match thirty words with their picture equivalents. The words and corresponding pictures were divided into three lists of four-, five-, or six-letter words. Each list contained ten words. The original design called for the presentation of the lists (in the above sequence) on three days per week, one per day. Each subject would receive the same list on the same day, and the sessions would cover 4 weeks with each list being presented once each week. However, unforeseen events, which will be discussed in a later section, prevented the precise execution of the planned design. In the procedure actually followed the order of the lists and the day of presentation were not the same for each subject. Every subject did receive four sessions of training on each list for a total of twelve sessions, as planned, and the training sessions were spaced over a period of 4 weeks.

Subjects

The subjects ($N = 36$) were educable mentally handicapped students (EMH) drawn from two intermediate public school classes located in the Cannady Elementary School, East St. Louis, Illinois. Each class of 25 students was taught by a team of two teachers. Ages of the students ranged from 9 years 4 months to 12 years 5 months. The general intellectual level of the pupils, as measured by the Wechsler Intelligence Scale for Children, ranged from 57 to 80.

All students were given the Word Recognition Subtest from the Botel Reading Inventory. On the basis of these scores they were divided into two blocks according to reading ability. The low ability group (LA) consisted of those students classified as non-readers, preprimer, and primer. The group of LA subjects had a mean CA of 10.7 and a mean IQ of 66.2. The high ability group (HA) contained pupils reading at the first, second, and third grade levels. The mean CA of the HA subjects was 10.9 and the mean IQ, 69.8. Within each block from each classroom three subjects were randomly assigned to each of three treatment groups. Therefore, the design consisted of 18 subjects from each classroom and 12 subjects in each treatment group. The independent variables were two levels of classrooms, two ability levels, and three treatments (see Figure 1).

Class-rooms	Reading Ability	Treatment Groups		
		I	II	III
X	LA	3	3	3
	HA	3	3	3
Y	LA	3	3	3
	HA	3	3	3
Number of Subjects		12	12	12

Fig. 1. Randomized blocks design with replication.

Apparatus

The teaching instrument (Figure 2) was constructed from a sheet of pegboard, 3½ feet wide and 2½ feet high. Attached to the board were 20 electrical contacts (screws). These contacts were arranged in two groups of two vertical rows, five per row. The ten contacts in the two rows on the left were connected by wires on the back of the board to contacts in the two rows on the right. These electrical connections were made in random sequence. A six volt battery was attached to the reverse of the board. Two probes on 4 foot leads were used to complete an electrical circuit lighting a bulb mounted in the top center of the board. Beside each of the contacts on the left-hand side of the board a picture of an object was placed. Under each picture was a separate and removable word which identified the picture. On the right-hand side of the board words corresponding to the pictures were arranged so that when the probes were touched to the correct picture-word contacts the electrical circuit for the light was completed signifying a correct response. The picture and word cards were affixed to the board with Plasti-Tak, a commercial adhesive. Use of this product permitted easy removal and switching of the cards.

Stimuli

As mentioned in the section on training, three lists of words, ten words per list, were constructed. They were; (List A) drum, pipe, lamp, fork, desk, moon, leaf, star, ring, tent; (List B) knife, truck, block, purse, glass, chair, canoe, spoon, fence,

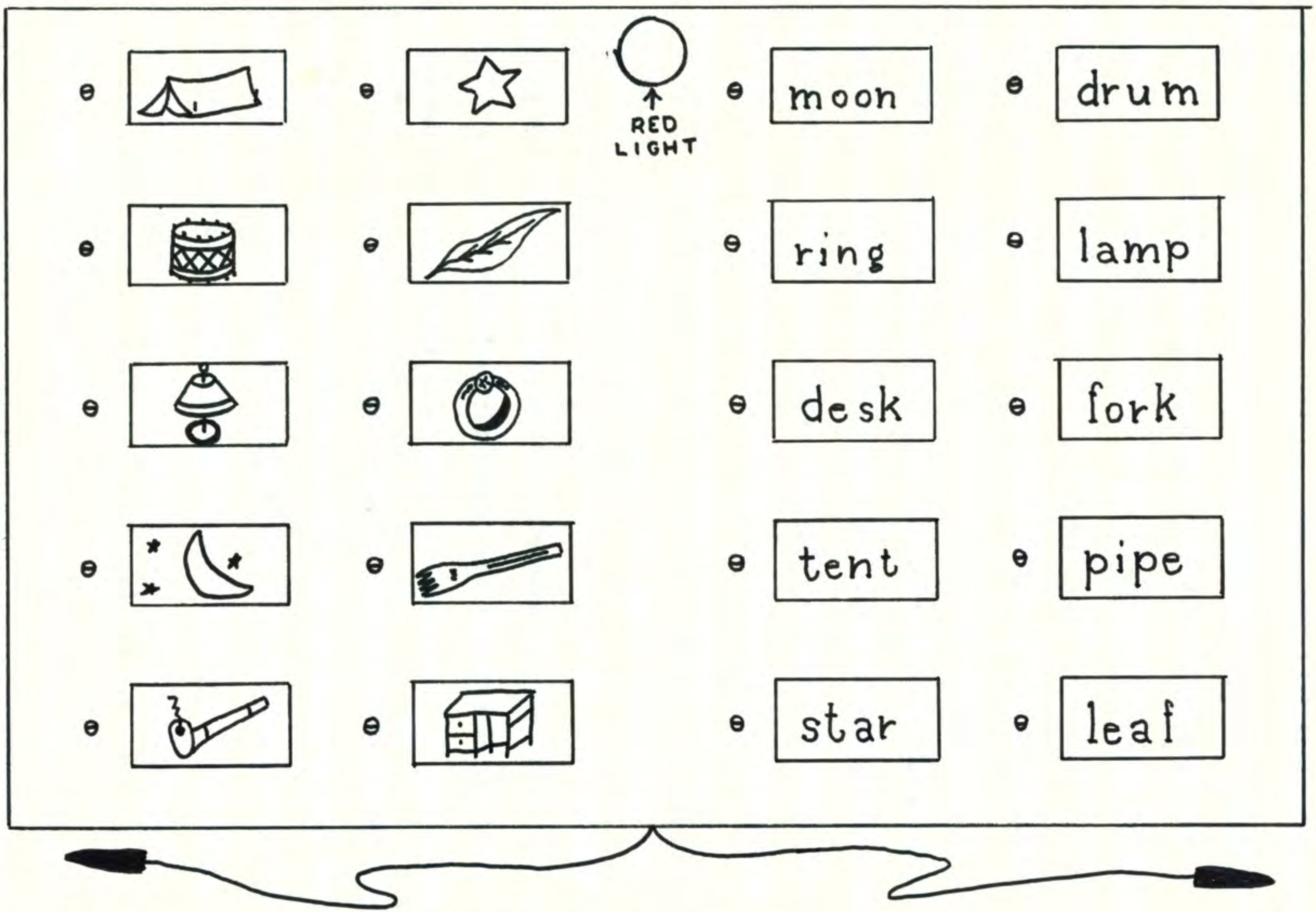


Fig. 2. Battery-Operated Teaching Instrument

ruler; (List C) rabbit, jacket, bottle, circle, square, pencil, window, hammer, banana, and candle. The criteria and steps involved in the choice of these words are recorded below.

1. All object-words of first-, second-, and third-grade reading categories were taken from The Teacher's Word Book of 30,000 Words (Thorndike & Lorge, 1944).

2. Words appearing in the basal readers of the EMH classes were deleted from the compiled list of Thorndike-Lorge words.

3. In addition, an effort was made to eliminate those words which could not easily be identified in picture form and words whose corresponding picture might have more than one familiar name.

4. Next, bigram frequencies for the pairs of letters in each word were added in order to give a frequency measure for the combination of letters in each word. The bigram frequencies (as found in printed matter) were obtained from Appendix D in Meaningfulness and Verbal Learning (Underwood & Schultz, 1960).

5. Words to be used in the study were selected on the basis of the totaled bigram frequency count for each word, so that a wide range of combinations of letters within words was represented. For example, List A contains "tent" with the highest count (8,954) and gradually descends to "drum" with the lowest count (900).

6. In the final selection an attempt was also made to prevent inter-list and intra-list similarity of both words and pictures.

Treatment Group I (Paper Group)

Twelve Ss (CA, 10.5; IQ, 69.4) serving as the control group remained in the classrooms for training. They received two papers during each session. Printed on the pages were ten pictures and ten words in two columns. The two papers for each list were identical except that one had an identifying word directly under each picture in addition to the randomly arranged column of words. The latter paper was given to the control Ss first. When it was completed and returned to the experimenter (E) they were given the second paper which did not have the additional cue words. Subjects were instructed to "Draw a line from the picture to the correct word." The papers were checked for mistakes and returned to the Ss the day of the next session. The group given this task served as a control group by representing "regular classroom practices" provided to these EMH students. In other words, the form and content of the papers was similar to tasks assigned to students as "seat-work".

Treatment Group II (Instrument Group)

Group II (n = 12; CA, 11.0, IQ, 68.9) received individual training on the teaching instrument. Training was conducted in cloakrooms adjoining the classrooms. One word list was presented at each session. A session consisted of eight trials; one trial was defined as correct responses to all pictures on the list, taken in the order appearing on the board. For four of the eight trials the correct word for each picture was placed below each picture giving the S a cue to the correct response. The cue words were removed from the board for the last four trials.

During the first session Ss were told that they were going to play a learning game, and, like all games, there were rules to be followed. One of the rules was that the E could not talk to the Ss while they were using the instrument. (Unfortunately, this rule was, of necessity, often broken for a variety of reasons.) The E did not indicate, either by gesture or verbally, approval or disapproval of a response.

Initially each S was given a demonstration. "Look, I'm going to touch a screw with this black pointer. Can you tell me the name of the picture next to this screw? That's right (or no, it's called -----). Watch what happens if I keep the black pointer where it is and place the red pointer on the screw beside the correct name for that picture. (The light came on.) Now you try to do what I did. Look for the big word that looks like the little word under the picture." If the S understood the directions he was told to proceed down the column, responding to each picture in order. If not, practice continued until the instructions were correctly followed.

On the first trial of every session each S was told to name the picture before pointing to the word.

Treatment Group III (Instrument Pair Group)

Group III (n = 12; CA, 10.7; IQ, 68.8) received the same instructions as Group II, but the Ss worked in pairs. For two of each four trials one S pointed to a picture and the other S made the response by placing the red probe on a screw next to a word. They were told that they could help each other find the correct word. Unless one S was absent the same two students were paired throughout the training sessions. Where possible, an absent student was replaced by another S from Group III (from the same classroom); if an odd number of absences occurred the E replaced the missing S by pretending to be a student.

Evaluation of the Data

Two reading tests (recall of words alone) were administered after completion of the twelve training sessions. The first test that was given contained the thirty words presented during the experiment. Data derived from this test were used to analyze the effects of the variables upon learning. The training words were printed in random order in two columns on an 8½ inch by 11 inch card. Ss were asked to read the words.

The second test consisted of the original thirty words plus thirty additional object-words which had not been presented during training sessions. The non-programmed words had the same number of letters as the training words and were also similar in grade-level reading difficulty. All words were listed in random order in three columns on an 8½ inch by 11 inch card.

The sixty word test was given for two reasons. First, in lieu of a pre-test, the data from the non-training words could be compared with that from the training words. Without special training all of the sixty words should have had an equal chance of being correctly identified, since none of the sixty words appeared in the basal readers of the participating classes. Therefore, a comparison of scores on the non-training words with scores on the training words furnished a measure of overall learning that might have occurred as a result of training.

The second reason for administering the test of combined words was to determine if ability to recall the training words would be affected with placement in a different context. It was reasoned that any one word in the list of thirty training words might be associated with the whole list. It would have been possible to correctly read a word by remembering the words that began with the same letter and guessing at the correct response. Combining the training words with non-training words in random ordered lists provided a method of assessing the effect of word placement on performance.

The original design called for testing the Ss one day after the last training session, and, again, seven days later in order to obtain retention scores. However, the high number of absences prevented the planned execution. Instead, all except two Ss in Classroom X were tested 3 days following the last session. The remaining Ss were given the tests 4 and 5 days, respectively, after training. Tests were administered to twelve Ss from Classroom Y five days following completion of the sessions. Three Ss received the tests on the eighth day and two Ss on the ninth day. One S was not tested.

CHAPTER III

Results

Assessment of Performance at End of Training

The thirty-word test: effects of treatments, classrooms, and reading ability. The mean for each block of *Ss* and the mean for the main effects are presented in Table 1. The data were taken from the thirty-word test given to the 36 *Ss* at the end of training.

TABLE 1

Block and Main Effect Means: Thirty-Word Test

	Means of Treatment Groups I	II	III	Reading Ability Means	Classroom Total Means
Class X					
Low Ability	4.00	9.00	6.33	6.44	15.66
High Ability	24.33	22.66	27.66	24.88	
Class Y					
Low Ability	8.00	9.66	7.00	8.22	16.78
High Ability	24.33	26.66	25.00	25.33	
Treatment Totals	15.17	17.00	16.50		
Low Ability Mean	7.33				
High Ability Mean	25.11				

Table 2 provides the data from an analysis of variance based on the thirty-word test. The score for the missing *S* was estimated using the method suggested by Kempthorne (1952, pp. 172-174). It is apparent from the table that the only significant main effect was Reading Ability. The highest performance on the criterion test was consistently made by Ss from the HA group. However, the relationship

TABLE 2

Analysis of Variance: Thirty-Word Test

Source	df	MS	F
Training Mode (T)	2	10.28	___ 1
Reading Ability (A)	1	2844.44	83.17*
Classrooms (C)	1	7.11	___
T X A	2	15.29	___
T X C	2	8.62	___
A X C	1	1.78	___
T X A X C	2	16.97	___
Within Treatments	24	34.20	___

1. (___) F value less than 1.00
* p less than .001

that seems to exist between previous achievement and ability to learn new words may simply represent a greater number of training words known by the HA Ss prior to the beginning of the study.

Absence of any significant differences between the three training procedures forces the objective conclusion that, within the context of this experiment, immediate feedback does not produce a higher level of reading performance in EMH children. The findings support the negative results of the studies of Blackman and Capobianco (1965), and Ellson et al. (1962), and are in partial agreement with Vergason's investigation (1967). In addition, paired training on the instrument is not superior to individual training, although it does not seem to have a detrimental effect upon final performance.

The sixty-word test: evaluation of the mean differences between training and non-training words. It may be recalled that a test consisting of the thirty training words plus thirty non-training words was given to the Ss to obtain a measure of generalized learning that might have occurred as a result of training. As shown in Table 3 the difference between means of training versus non-training words is significant at or beyond the .025 level (22 df for all groups, n = 12). This suggests that some degree of learning occurred, and that it did so under all three treatment conditions. However, in the absence of a pre-test or a control group who received no special training, the conclusion is somewhat tentative.

TABLE 3

A Comparison of Recall Scores Between Training

and Non-Training Words

Treatment Condition	Means: T Words	Means: Non-T Words	LA & HA: T Words	LA & HA: Non-T Words	t
LA Group I	5.2	.5			
HA Group I	23.0	10.7	14.1	5.6	2.19*
LA Group II	9.8	1.3			
HA Group II	24.7	13.0	17.3	7.2	2.65**
LA Group III	6.9	1.3			
HA Group III	25.7	12.3	16.3	6.8	2.60**

*p less than .025 (22 df)
**p less than .01 (22 df)

As previously indicated, the other reason for administering the test of combined words was to determine if ability to recall the training words would be affected with placement in a different context. In fact, inability to recall a word on one test while correctly identifying it on another test occasionally happened. This type of error was distributed over 23 of the 30 words. As seen by Table 4, with means comparable between tests, neither test showed superior performance, but lack of perfect reliability in performance was evident for 24 Ss. Nonetheless, correlation was .99 across all Ss which gives an overall measure of reliability of their reading performance.

TABLE 4

Comparison of Recall of Training Words

on Thirty-Word and Sixty-Word Tests

	LA Group	HA Group	LA & HA Groups
No. of words correctly recalled on 30-word test but not on 60-word test	11.	19.	30.
No. of words correctly recalled on 60-word test but not on 30-word test	11.	8.	19.
No. of Ss making one or both kinds of error	10.	14.	24.
Mean no. of words correctly recalled on 30-word test	7.33	25.11	16.22
Mean no. of T words correctly recalled on 60-word test	7.33	24.50	15.92
Correlation of T word scores on both tests	.97	.96	.99

Bigram frequencies and word recall. It will be recalled that the training words were selected partly on the basis of the totaled bigram frequencies for each word. In order to determine if the bigram count (as found in printed material) was related to reading difficulty for EMH children data from the Thirty-Word Recall Test and the bigram count for each word were compared. The statistical finding was a non-significant correlation coeffient of -.14.

Since number of letters might affect reading performance, recall scores of the four-, five-, and six-letter word lists were compared. A Kruskal-Wallis analysis revealed a Chi^2 value of 3.0 which was not significant.

Assessment of Performance During Training

The effect of errors on performance. During the training sessions the number of errors made by each S was recorded. Preliminary appraisal of the data indicated that the number of errors between groups might be significantly different. However, statistical analysis was somewhat complicated because of two factors. First, one S from each instrument training group had very extreme scores which negated the use of an analysis of variance. Secondly, and more importantly, the analysis involved a comparison of errors made under essentially different procedures. As indicated in Table 5 the maximum number of errors that could be made under each treatment condition differed greatly. Therefore, in order to have an understanding of training errors, computation of the data was conducted in several ways. In all cases, however, a Kruskal Wallis "analysis of variance" by ranks was performed. A comparison of the results using different methods of computing the errors is presented in Table 5.

TABLE 5

Comparison of Errors

	Treatments I	II	III
Median no. of errors over 12 sessions	9.5	32.5	27.5
Range of errors over 12 sessions	0-79	1-874	1-164
Max. no. of possible errors for one training session	20	720	360
Max. no. of possible errors over 12 sessions	240	8640	4320

Computation of Errors	Rank-Sums of Treatment Conditions I	II	III	Chi^2
Percentage of max. errors over all trials	329.0	151.5	185.5	13.3*
Total errors over all trials	181.0	246.5	238.5	1.9
First error on each word over all trials	195.5	247.0	223.5	.1
First error on each word over two trials in each session	244.0	193.4	227.5	1.0

Note.__ Higher rank-sums indicate a greater number of errors.
*p less than .01

The first procedure involved a conversion of the raw scores into percentages of the maximum number of possible errors for each group. When errors are computed in this manner the Paper Group (Group I), who made the lowest number of actual errors, obtained the highest rank-sum. It will be noted from Table 5 that this measure produced the only significant result. The second analysis involved a comparison of the total number of errors over all 12 training sessions. In this case, however, the raw-error score of each S was used (the scores were not converted to percentages of the maximum number of possible errors for each group). Thirdly, a comparison was made of the number of errors on the first response to each word over all trials. The scoring was conducted on a correct or incorrect basis with all subsequent incorrect responses to a word on the same trial being disregarded. Since Group I made only one response to a word on each trial, the scores for this group were not changed. However, seven error-scores in each of the instrument training groups were reduced. For the final analysis error-scores were computed on the first trial of cue training and the first trial of non-cue training for the 12 sessions. Group I was given just two trials (two papers) during each session, so the error scores for this group remained unchanged.

The analysis of errors shows that the rank-sums of Group II and III were only slightly different under the four computational procedures. The situation is interesting since the opportunity for the Instrument Pairs to make errors was half as much as that for Ss working individually. A possible explanation may be found in the "help" that was given the active S by the observing partner. Perhaps the "help" frequently led the responder to make the wrong response as well as the correct response. It could be postulated, then, that an error-score represented the errors of not one, but two Ss over four trials. In effect, this would equal the score of one S over eight trials.

Although the only significant result occurred when the errors were transformed to percentages, Group I also had the highest rank-sum when the number of trials for each group was matched. If viewed from a relative, rather than absolute position, the performance of Group I during training was inferior to that of the Instrument Groups. Indeed, considering the number of errors that the Instrument Ss could have made, their performance was rather surprising.

As previously noted, the greatest number of actual errors was made under the instrument training procedures, but it is evident from Table 6 that error differences were confined mainly to the LA Groups. During the first training session for each word list LA Ss in the Instrument Groups made a far greater number of errors than LA Ss in the Paper Group. By the final session of each list, however, median errors of the LA Instrument Groups had decreased to the extent that generally the errors equaled those made by the LA Paper Group. In contrast, errors were comparable over sessions for HA Ss in the Instrument Groups and Paper Group.

TABLE 6

Median Errors Over Treatment Groups For Each List

Word List	Training Session	Reading Ability	Median Errors Over Treatment Groups I	II	III
Four Letter List	1	LA	2.0	3.0	8.0
		HA	0.0	.5	0.0
	2	LA	2.5	9.5	5.0
		HA	0.0	1.0	0.0
	3	LA	0.0	1.0	6.0
		HA	0.0	0.0	0.0
	4	LA	.5	1.0	0.0
		HA	0.0	0.0	0.0
Five Letter List	1	LA	3.0	10.0	10.0
		HA	0.0	.5	0.0
	2	LA	3.5	3.5	1.5
		HA	0.0	.5	0.0
	3	LA	2.0	2.5	4.5
		HA	0.0	0.0	.5
	4	LA	1.0	.5	1.0
		HA	0.0	0.0	0.0
Six Letter	1	LA	4.0	8.5	5.5
		HA	2.0	1.0	2.0
	2	LA	4.5	6.5	1.5
		HA	0.0	.5	1.5
	3	LA	5.5	3.0	3.5
		HA	0.0	.5	1.5
	4	LA	2.5	4.5	2.0
		HA	.5	0.0	2.0

CHAPTER IV

Discussion

The Experimental Setting

Conducting research in a natural environment, such as a school, can be particularly difficult because of unexpected and uncontrollable events. The present investigation was no exception. It may be recalled that parts of the original design had to be eliminated or amended. Existing conditions, which will be described, prevented controlled sequencing of training sessions and the administration of retention tests, and made it necessary to reduce the number of trials included in the initial plan. The investigation, including preliminary testing, was to extend through the last two months of the school year. It was delayed two weeks, however, because of a teachers' strike and because many students were absent after the resumption of school. Permission to conduct the study was obtained from the classroom teachers on the condition that Ss from one classroom would be trained during the morning and Ss from the other classroom would be trained in the afternoon. Since the morning session was three hours long and the afternoon only two, this arrangement created a time differential which was compounded by other factors. The teachers in the "morning" classroom were very flexible and encouraged the use of the students at any time. The "afternoon" teachers, however, were somewhat apprehensive during the first half of the study and limited training sessions to the periods when the Ss were not occupied with other activities. Also, as it happened, two Ss in the "afternoon" Instrument Group were very slow, taking two or three times longer than was required for the other Ss. Because of these unexpected events the lists of words were not presented to each S on the same day or in the same sequence. It is doubtful that sequencing of training sessions was important in determining the outcome of the experiment. However, insufficient time for training sessions (in the "afternoon" class) made it necessary to reduce the number of trials originally perceived as minimum for an adequate test of the problem, and number of trials may have been a significant factor.

Irregular school attendance also prevented administration of retention tests at one and seven day intervals which was included in the original design. Recall Tests could not even be administered to all Ss within the same number of days after completion of training. Evaluation of data with regard to retention was, therefore, impossible. Retention may be a crucial element since Vergason (1967) found no difference between groups after one day, but significant differences at one and fourteen months in favor of auto-instruction. However, Blackman & Capobianco (1965) reported no significant differences between machine and "no-machine" groups at one day or 3½ months. Opposite results of the two experiments may be due to methodological variation.

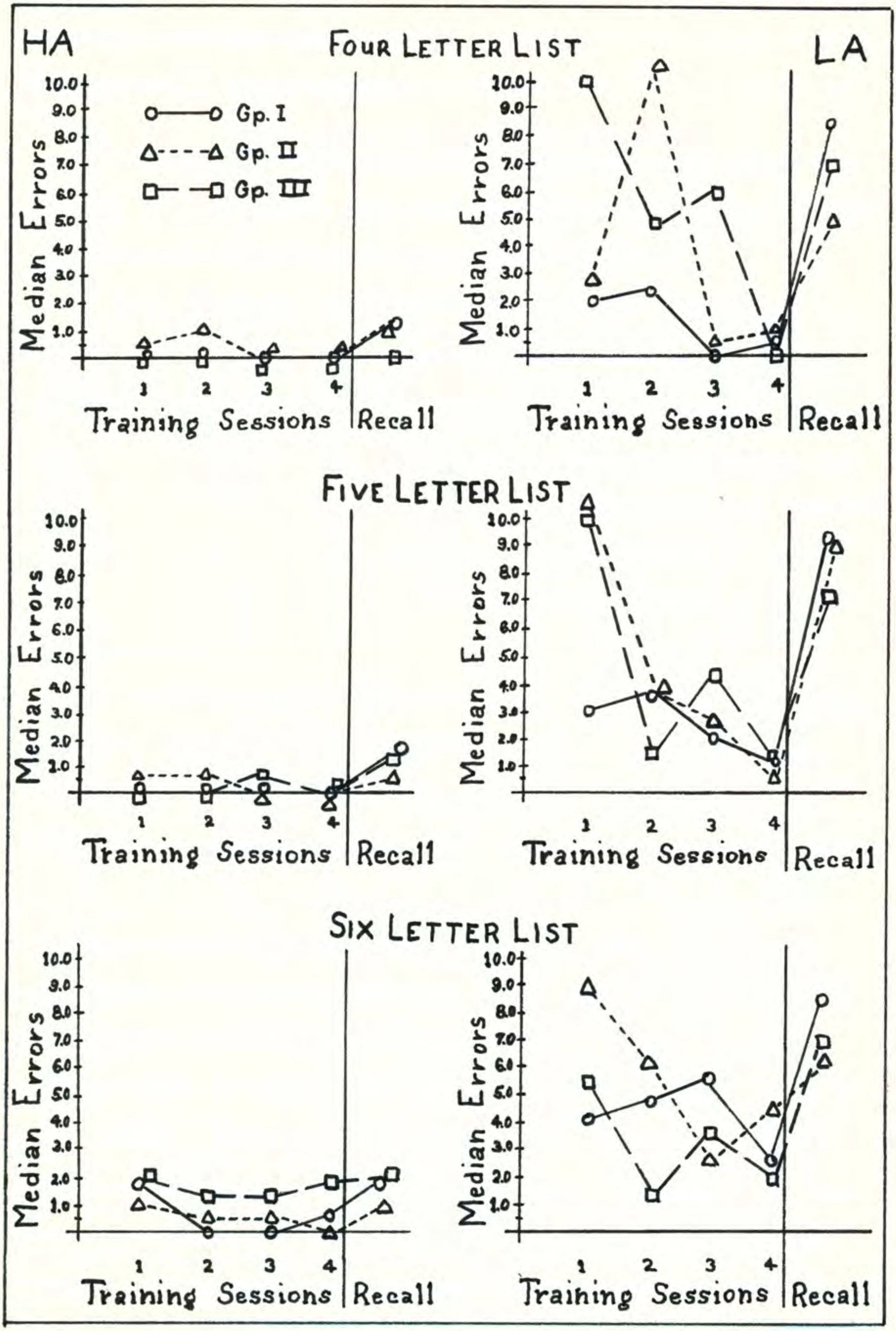
Analysis of Results

The significant difference between the HA Group and LA Group on the Thirty-Word Recall Test reflects, of course, a greater number of words recalled by the HA Group. Other investigators

(Ellson, Engle, & Barber, 1963; Malpass et al., 1963) have also noted that those Ss who had higher mean scores before instruction demonstrated higher mean scores after instruction. At least two possible explanations for the higher performance of the HA Ss can be offered: (1) Learning to read new words becomes progressively easier as the sight vocabulary increases, and (2) the relationship that seems to exist between previous achievement and ability to learn new words may simply represent a greater number of training words known by the HA Ss prior to the beginning of the experiment. Although the first explanation cannot be discounted, evidence from the present study, at least, suggests that the second explanation is more tenable. LA Ss over all treatment conditions correctly read an average of 3% of the non-programmed words and 24% of the programmed words, or 8 times as many. HA Ss identified 40% of the non-programmed words and 82% of the programmed words, or just twice as many. The apparent reason for the discrepancy lies in the low ceiling of the task for the HA Ss. With lists of constant length the number of new words that the HA Ss could learn was relatively (though not actually) less when compared to the LA Group. Confirmation of the low ceiling for HA Ss was found in the analysis of performance during training. It may be recalled that there was an absence of error differences for HA Ss in contrast to the result for the LA Groups. In addition, the slopes of the learning curves for the HA Groups (Figure 3) were almost rectilinear indicating there was little, if any, improvement with practice.

Conversely, the task was more difficult for the LA Ss, and errors decreased over sessions showing improvement in performance with practice. Furthermore, the slopes were steeper for the LA Instrument Groups than for the LA Paper Group, and the LA Ss identified a slightly greater number of words on each list of training words than did the Ss from the Paper Group. The results suggest that additional trials or sessions might have increased the performance of the Instrument Groups (LA) to a significant level.

Fig. 3. Median Training Errors and Median Recall Errors over Treatment Groups for each Word List.


HA
FOUR LETTER LIST
LA
Gp. I
Gp. II
Gp. III
Median Errors
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
1
2
3
4
Training Sessions
Recall
FIVE LETTER LIST
SIX LETTER LIST

Methodological Inadequacies and Suggestions for Future Research

In addition to deficiencies which have already been discussed, a number of other factors have been identified which may have affected any conclusions drawn from the experimental operation.

A major defect was the omission of a pre-test. Comparison of scores on the training and non-training words showed a significant difference favoring the training words. The result indicates that training was effective under all treatment conditions. However, the actual gain in word recall is unknown. A pre-test would have provided comparison scores. It would also have been advantageous in determining the number of training words known by the Ss prior to the investigation.

In addition, research with normals and a few studies with mental retardates suggest that the conditions of task presentation during training may not have been the most beneficial for learning. One aspect of the experiment which probably affected the rate of learning was the picture-word presentation of the items. It has been found in paired-associate learning with normals that familiarization and meaningfulness of response items has a more important function in the facilitation of learning than familiarization and meaningfulness of stimulus items (Morikawa, 1959; Underwood & Schultz, 1960). A recent study (Hawker, 1968) on learning of sight vocabulary by mental retardates confirms the results of experiments with normals. In the Hawker investigation a Recall Test after paired-associate training revealed that the word-picture method of presentation was significantly superior to the picture-word method. Training procedures were different from the present study, however, in that only one stimulus item and 4 alternative responses were presented at one time. Additional research is needed to determine if the same result would occur when a larger number of stimulus and response items appear simultaneously.

Interacting with the picture-word presentation, and possibly reducing performance during both training and recall was the absence of a required oral response during training. It may be recalled that the Ss verbally identified each picture on the first trial of every session, but were not required to give an oral response to the word. If, at the end of training, a recognition test had been given, in which the task consisted of matching a word with the correct picture, the procedure would have produced stimulus-response learning during acquisition, but backward (R-S) recall. A series of experiments with normals conducted by Morikawa (1959) established that backward recall was inferior to forward recall. However, the final test of learning was not an association of a word with a picture, (A-B, B-A paradigm), but an oral response to a visual stimulus (A-B, B-C paradigm). In effect, the recall task called for a mediating response which was not included in the acquisition procedure. Therefore, the question arises, "Did the Ss produce the mediating verbal response during training, (that is, did they label the pictures and words) as normals would? Or was the association a purely visual phenomenon?" A number of experiments have shown that verbalization facilitates discriminative or associative learning in the mentally retarded (Barnett, Ellis & Pryer, 1959; Jensen & Rohwer, 1963; Berkson & Cantor, 1960;

Griffith & Spitz, 1958; Miller & Griffin, 1961). However, the majority of these studies used Trainable Ss rather than EMH, and none of the investigations specifically tested the incorporation of verbal mediators when the items to be associated were a meaningful pictured object and the identifying word. Of the four studies previously described which compared programmed instruction with traditional classroom teaching, two required an oral response during training and two did not. The results did not favor either response method. If future investigations show that EMH children do not automatically supply the necessary verbal mediators, an interesting set of problems is presented with regard to self-instruction. (1) If initial instructions include the admonition to respond verbally to each item, will the Ss continue to do so on their own initiative? (2) How useful is the "seatwork" which is assigned as practice lessons (especially for the beginning EMH reader)? Should students be encouraged to read aloud when working alone? (3) Can verbal responses be programmed on teaching instruments so that the machine will indicate a correct or incorrect response? A few highly sophisticated (and expensive) computers can react to an auditory input, but speech defects often encountered in the mentally retarded would possibly present insurmountable technical problems.

Unfortunately, the present experiment did not adequately test the effects of immediate feedback on learning as compared to conventional classroom procedures. The study did, however, direct attention to the problems of conducting research in a public school located in a deprived area. Absenteeism, for instance, can destroy the most carefully designed experiment. Therefore, the attendance rate should be a prime concern when planning a study. Perhaps school records would show a trend in attendance indicating that some months would be superior to others with regard to research.

This study also points to the great need for investigation of the learning processes of the mentally retarded. Relatively little research has been conducted with non-institutionalized EMH children so that studies in this area could contribute much of practical and theoretical importance.

CHAPTER V

Summary

A comparison was made between two methods of teaching sight vocabulary to educable mentally retarded children placed in special public school classes. Thirty-six Ss trained on a task consisting of learning three lists of words (10 items per list) which were paired with their picture counterpart. A 3 x 2 x 2 factorial design was employed in which the three variables were training procedure (Paper, Instrument, or Paired Instrument), classroom, and ability level (High or Low Ability). A Recall Test for words only was given after completion of four sessions of training on each list.

Evaluation of the data revealed that training procedures and classrooms were not significant factors in final performance. There was a significant difference between the mean recall scores of the High Ability and Low Ability Ss with the High Ability Group making consistently higher scores after completion of training.

A comparison of scores made on an additional test consisting of the training words and similar non-training words resulted in a significant difference favoring the training words. In absence of a pre-test this was an indication that training improved the performance of all groups of Ss.

Methodological inadequacies of the study were discussed which might have affected any conclusions, and suggestions were made for future research comparing immediate knowledge of results and individual student assignments ("seatwork") using mentally retarded children as subjects.

Fig. 4. Example of Paper Given to Group I (Paper Group)

TABLE 7

Chronological Age and IQ of Subjects

Class-room	Reading Ability	Subject	CA (months)	IQ
X	HA	B.R.	135	72
		R.D.	135	72
		J.G.	134	73
		D.W.	134	66
		K.S.	134	73
		A.H.	137	73
		C.R.	135	78
		M.H.	138	63
		D.S.	135	70
X	LA	C.T.	136	68
		L.R.	131	60
		M.W.	130	72
		K.S.	136	68
		D.D.	144	62
		R.S.	143	73
		L.P.	130	70
		D.D.	137	56
		M.B.	139	65
Y	HA	C.W.	120	64
		T.T.	115	65
		D.J.	128	77
		M.P.	129	65
		F.J.	119	73
		M.S.	136	65
		E.J.	122	80
		C.S.	122	67
		W.S.	126	62
Y	LA	C.B.	116	60
		P.E.	116	80
		M.M.	112	60
		M.B.	118	61
		J.P.	125	67
		J.C.	131	67
		B.M.	116	70
		S.T.	126	57
		P.D.	114	76

TABLE 8

Totaled Bigram Frequencies

of Training Words

Word	Totaled Bigram Frequency
Drum	900
Pipe	1778
Lamp	2448
Fork	3220
Desk	3809
Moon	4775
Leaf	5485
Star	6775
Ring	7415
Tent	8954
Knife	1887
Truck	1949
Block	3055
Purse	3859
Glass	4076
Chair	4583
Canoe	5058
Spoon	5348
Fence	6306
Ruler	7785
Rabbit	3715
Jacket	4073
Bottle	4993
Circle	5621
Square	6793
Pencil	6947
Window	6973
Hammer	7402
Banana	8051
Candle	8831

REFERENCES

Amaria, R. P. Learning in "pairs". Proceedings of the APLNCPL Programmed Learning Conference, April 1967. Cited by J. Hartley, Some factors affecting student performance in programmed learning. Programmed Learning and Educational Technology, 1968, 5, 206-218.

Austwick, K. "Research report". New Education, 1965, 1, (10), 25-26. Cited by J. Hartley, Some factors affecting student performance in programmed learning. Programmed Learning and Educational Technology, 1968, 5, 206-218.

Amaria, R. P., Biran, L. A., & Leith, G. O. M. Learning in "pairs". Research Notes on Programmed Learning, 1966, No. 10, National Center for Programmed Learning, University of Birmingham, England. Cited by J. Hartley, Some factors affecting student performance in programmed learning. Programmed Learning and Educational Technology, 1968, 5, 206-218.

Barnett, C. D., Ellis, N. R., & Pryer, N. W. Stimulus pre-training and delayed reaction in defectives. American Journal of Mental Deficiency, 1959, 64, 104-111.

Bennett, A. A comparative study of sub-normal children in the elementary grades. New York: Teachers College, Columbia University, Bureau of Publications, 1932. Cited by S. A. Kirk, Research in education. In Stevens, H. A., & Heber, R. (Ed.), Mental retardation: A review of research. (2nd ed.) Chicago, University of Chicago Press, 1965. Pp. 57-99.

Berkson, G., & Cantor, G. N. A study of mediation in mentally retarded and normal school children. Journal of Educational Psychology, 1960, 51, 82-86.

Blackman, L. S., & Capobianco, R. J. An evaluation of programmed instruction with the mentally retarded utilizing teaching machines. American Journal of Mental Deficiency, 1965, 70, 262-269.

Blatt, B. The physical, personality, and academic status of children who are mentally retarded attending special classes as compared with children who are mentally retarded attending regular classes. American Journal of Mental Deficiency, 1958, 62, 810-818.

Cassidy, V. M., & Stanton, J. E. An investigation of factors involved in the educational placement of mentally retarded children: A study of differences between children in special and regular classes in Ohio. (U. S. Office of Education Cooperative Research Program," Project No. 043.) Columbus: Ohio State University, 1959. (Washington, D. C., Library of Congress).

Dick, W. Retention as a function of paired and individual use of programmed instruction. Journal of Programmed Instruction, 1963, 2(3), 17-23.

Eigen, L. D. Problems of research in programmed instruction. A-V Communication Review, 1965, 13, 38-43.

Ellson, D. G., Engle, T. L., Barber, L., & Kempworth, L. Programmed teaching of elementary reading--a progress report. Unpublished manuscript, Indiana University, 1962. (Supported by USPHS grants M4786 (A) and M4989.) Cited by Francis M. Greene, Programmed instruction techniques. In N. R. Ellis (Ed.), International review of research in mental retardation. New York: Academic Press, Vol. 1, 1966. Pp. 209-240.

Goldstein, H., Jordan, L., & Moss, J. W., The efficacy of special class training on the development of mentally retarded children. (Cooperative Research Program of the Office of Education, U. S. Department of "HEW," Project No. 619.) Champaign: University of Illinois 1965. (Library of Congress, Washington, D. C.)

Griffith, B. C., Spitz, H. H., & Lipman, R. E. Verbal mediation and concept formation in retarded and normal subjects. Journal of Experimental Psychology, 1959, 58, 247-250.

Kaess, W., & Zeaman, D. Positive and negative knowledge of results on a Pressey-type punchboard. Journal of Experimental Psychology, 1960, 60, 12-17.

Kempthorne, O. The design and analysis of experiments. New York: Wiley, 1952.

Kirk, S. A. Educating exceptional children. Boston: Houghton Mifflin, 1962.

Knight, S. Programming for the retarded reader. Programmed Instruction, 1963-1964, 3, 3+.

Malpass, L. F., Hardy, M. W., Gilmore, A. S., & Williams, C. F. Automated instruction for retarded children. American Journal of Mental Deficiency, 1964, 69, 405-412.

Miller, M. B., & Griffith, B. C. The effects of training verbal associates on the performance of a conceptual task. American Journal of Mental Deficiency, 1961, 66, 270-276.

Morikawa, Y. Functions of stimulus and response in paired-associate verbal learning. Psychologia, 1959, 2, 41-56.

Mullen, F. A., & Itken, W. Achievement and adjustment of educable mentally handicapped children. ("Cooperative Research Program of the Office of Education, United States Department of Health, Education, and Welfare," Project No. SAE 6529.) Chicago: Board of Education, City of Chicago, 1961. (Library of Congress, Washington D. C.)

Peterson, L. R., & Brewer, C. L. Confirmation, correction, and contiguity. Journal of Verbal Learning and Verbal Behavior, 1963, 1, 365-371.

Russell, D. H., & Fea, H. R. Research on teaching reading. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand-McNally, 1963. Pp. 865-928.

Sawaris, M. An experimental study of individual and group learning using a linear geometry program. Programmed Learning, 1966a, 3(3), 146-153.

Stolurow, L. M. Implications of current research and future trends. Journal of Educational Research, 1962, 55, 519-527.

Thorndike, E. L., & Lorge, I. The teacher's word book of 30,000 words. New York: Teachers College, Columbia University, 1944.

Underwood, B. J., & Schultz, R. W. Meaningfulness and verbal learning. Chicago: Lippincott, 1960.

Underwood, B. J., & Schultz, R. W. Studies of distributed practice: XX. Sources of interference associated with differences in learning and retention. Journal of Experimental Psychology, 1961, 61, 283-289.

Vergason, G. A. Retention in educable retarded subjects for two methods of instruction. _American Journal of Mental Deficiency_, 1966, 70, 683-688.

Warren, K. _An investigation of the effectiveness of educational placement of mentally retarded children in a special class._ (Doctoral dissertation, University of Tennessee) Ann Arbor, Mich.: University Microfilms, 1962. No. 59-297.

Wrightstone, J. W., Forlano, G., Lepkowski, J. R., Sontag, M., & Edelstein, J. K. _A comparison of educational outcomes under single-track and two-track plans for educable mentally retarded children._ ("Cooperative Research Program of the Office of Education, United States Department of Health, Education, and Welfare," Project No. 144). New York: New York Board of Education, 1959. (Library of Congress, Washington, D. C.)

Graduate School
Southern Illinois University

me: Betty Jean (Wolter) Wilson

Local Address: 7 Hamlet Court
Belleville, Illinois

Home Address: Same

lleville Township Junior College	1948-50	A.A.	Liberal Arts
uthern Illinois University	1963-66	B.S. in Ed.	Psychology
uthern Illinois University	1966-69	M.S.	Psychology

mber of Psi Chi

esis Title: Three Methods of Teaching Sight Vocabulary to
Educable Mentally Handicapped Children

viser: Dr. Eva D. Ferguson

Graduate School
Southern Illinois University

Name: Betty Jean (Wolter) Wilson

Local Address: 7 Hamlet Court
Belleville, Illinois

Home Address: Same

Belleville Township Junior College	1948-50	A.A.	Liberal
Southern Illinois University	1963-66	B.S. in Ed.	Psychol
Southern Illinois University	1966-69	M.S.	Psychol

Member of Psi Chi

Thesis Title: Three Methods of Teaching Sight Vocabulary to Educable Mentally Handicapped Children

Adviser: Dr. Eva D. Ferguson