DOCUMENT RESUME

ED 026 751 EC 001 270

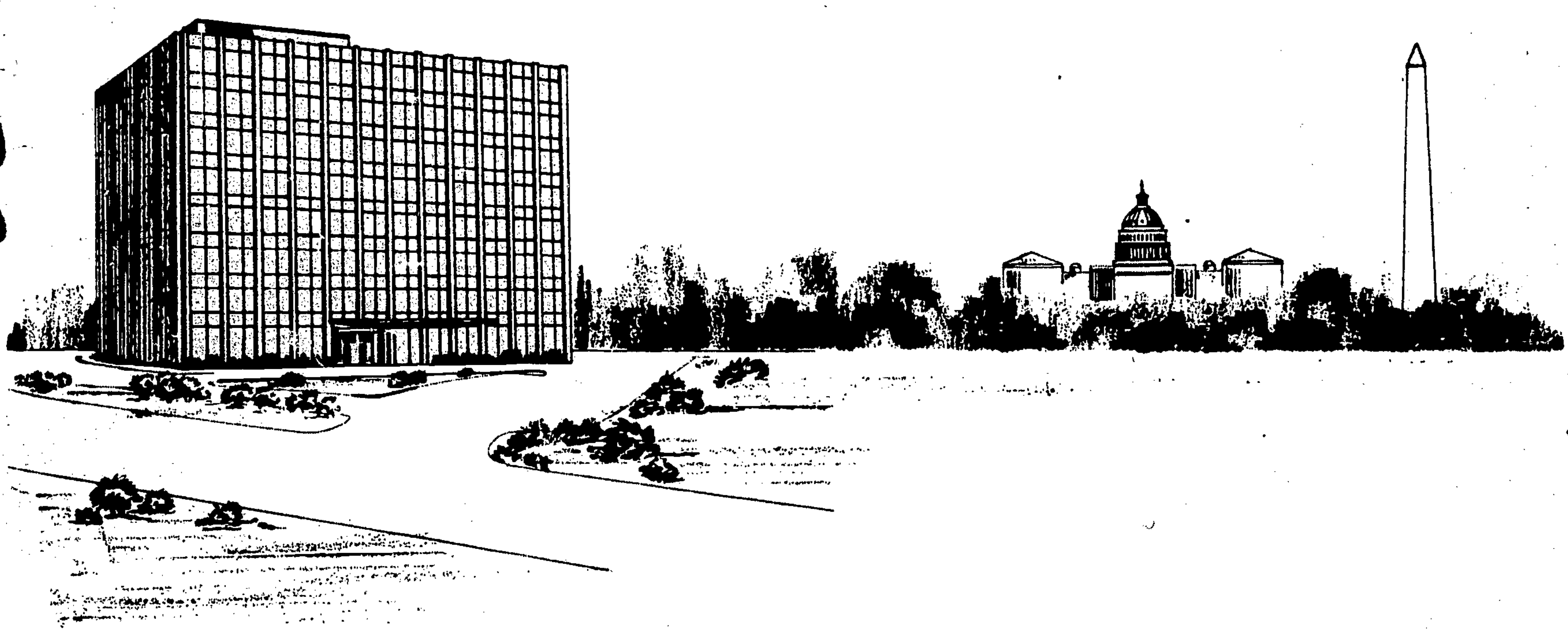
By-Robinson, Jacques H.; And Others
Bimodal Educational Inputs to Educable Mentally Retarded Children. Final Report.
American Institutes for Research, (Washington Office) Silver Spring, Md. Communication Research Program.
Spons Agency-National Inst. of Health, Bethesda, Md.
Report No-AIR-E-49-9-66-FR
Pub Date Sep 66
Grant-MH-10819-01
Note-147p.
EDRS Price MF-$0.75 HC-$7.45
Descriptors-Audio Equipment, *Audiovisual Instruction, Aural Learning, Educable Mentally Handicapped, *Exceptional Child Research, Instructional Technology, Intermode Differences, *Learning, Listening Comprehension, *Mentally Handicapped, Multimedia Instruction, Multisensory Learning, Pacing, Reading Comprehension, Retention, Speech Compression, Tape Recordings, Teaching Methods, Visual Learning

The usefulness of paced auditory presentation combined with simultaneous visual presentation of lesson material was examined as a technique of improving reading skills in the educable mentally handicapped. A 30-day instructional period involved four groups of students (IQ's 58 to 86, ages 12-11 to 17-11). The groups were as follow: a machine audio group (MAud) exposed to rate controlled taped material by audio means alone; a machine audiovisual group (MAV) that received visual copy and also listened to the taped recordings; a teacher audiovisual group (TAV) using bimodal stimulation; and a control group (C). Materials were recorded and presented at appropriate rates of presentation by means of a tempo regulator, a machine which varies rate without pitch distortion. Results indicated little impact of the experimental procedures on the standardized pre-post test measures; however, daily comprehension measures favored the bimodal presentation modes, particularly the MAV group. Some trend favoring slowed presentations was evident. The subjects demonstrated significant retention of instructional material after a 1-month interval, and significantly higher performance on a relearning measure than on initial presentations, after a 2-month interval. (Author/JD)

Bimodal Educational Inputs to Educable Mentally Retarded Children

Jacques H. Robinson
David B. Orr
Victor K. Small

Final Report
SEPTEMBER 1966



AMERICAN INSTITUTES FOR RESEARCH
WASHINGTON OFFICE
Address: 8555 Sixteenth Street, Silver Spring, Maryland 20910
Telephone: (301) 587-8201

BIMODAL EDUCATIONAL INPUTS TO EDUCABLE MENTALLY RETARDED CHILDREN

FINAL REPORT

Grant No.: MH 10819-01

National Institutes of Health
Bethesda, Maryland

by

Jacques H. Robinson
Project Director

David B. Orr
Principal Investigator

Victor H. Small

American Institutes for Research

Washington Office

Communication Research Program

September 1966

AIR-E-49-9/66-FR

Acknowledgments

The authors wish to express their sincere appreciation to the many persons whose cooperation made the study possible:

The students and staff of the Mary Zigler School and personnel of the Laurel District (D. C.) Childrens Center for participation and assistance in conduct of Experiment I.

The students and staff of Rock Terrace High School and personnel of the Montgomery County (Md.) Public Schools and Board of Education for participation and assistance in the conduct of Experiment II.

The Bobbs-Merrill Publishing Company for permission to reproduce portions of the Pressey Diagnostic Reading Test.

The Lyons & Carnahan Publishing Company for discount on purchase (for research purposes) of the Bond-Clymer-Hoyt Diagnostic Reading Test.

The Educational Developmental Laboratories for supplying equipment and materials used for research purposes.

Those members of the AIR staff who contributed to the preparation of materials; Ruth Finney, John D. Zimmerman, and Dr. Mary Willis; the secretarial and reproduction staff; and most particularly to the Project Assistant, Mary Ann Brice, for her invaluable help.

Abstract

This project examined the usefulness of paced auditory presentation combined with simultaneous visual presentation of lesson material as a technique of improving reading skills in EMRs. A 30-day instructional period involved four groups of students enrolled in a special education school of the Montgomery County School System, and included: a group of EMRs exposed to rate controlled taped material by audio means alone (MAud); a group that received visual copy and also listened to the taped recordings (MAV); a group taught by qualified teachers using bimodal stimulation (TAV); and a control group (C) that did not participate in the experimental program. Selected and edited materials were recorded and presented at appropriate rates of presentation by means of a Tempo Regulator, a machine which varies rate of presentation by recorded material without pitch distortion. Standardized tests were administered before and after the period and each daily session was followed by a short comprehension test measure. Progress in reading and listening comprehension among the groups was compared.

The results indicated little impact of the experimental procedures on the pre-post measures; however, daily comprehension measures favored the bimodal presentation modes, particularly the MAV group. Some trend favoring slowed presentations was evident. EMRs demonstrated significant retention of instructional material after a one-month interval, and significantly higher performance on a re-learning measure than on initial presentations, after a two-month interval. It was concluded that the results were sufficiently promising to warrant further experimentation along these lines.

Table of Contents

Introduction

The research described in this report is part of a program of research which has as its long-term goal the investigation and improved understanding of human informational inputs, processing, and storage, and the translation of such improved understanding into both theory and practice. The research described here had two broad aims: (a) to study the effect of using simultaneous, bimodal (eye-ear) inputs on the comprehension of connected discourse, both in terms of assimilation and retention of content and in terms of improvement of reading and listening skills; and (b) to consider the implications of such procedures for the instruction of that portion of the intellectual continuum commonly known as educable mentally retarded (EMR).

As background, three specific earlier studies relevant to the present research are discussed below. The variables explored in these three studies were of concern in the present study and the findings formed an important basis both for the design of specific procedures and for the interpretation of the findings in the present instance.

The Heckelman Study

In addition to work conducted by the present Principal Investigator (Orr, et al, 1965), a study by Heckelman (1964) served as the immediate stimulus for the present research effort. Heckelman reported the use of simultaneous oral reading by teacher and student as a remedial reading training procedure: "Children with severe handicaps in the area of reading were given a maximum of 7-1/4 hours of instruction by this method during a period of six weeks with a resultant average of 2.2 grade levels of growth in functional reading skill." This kind of growth as a function of so little instruction is nothing short of phenomenal, and demands follow-up and replication.

If the gain in the Heckelman study can be attributed to either a Hawthorne or a Crespi-Zeaman "elation" effect, the increment relates to a performance rather than a learning variable. In practical terms, it may be sufficient that the growth obtains regardless of why it comes about. On the other hand, it is important to determine the relative role of associative and motivational factors in such gain in order to design optimally effective and efficient teaching techniques.

The present study may be considered to replicate some aspects of Heckelman's study. There are, however, salient differences between the two studies. The present research:

(a) was conducted with EMR subjects.

(b) employed group instructional methods.

The use of group instruction precluded the possibility of having the children read aloud. This, in turn, precluded using such oral reading to insure that the pupils are attending to the task. Simultaneous oral reading on the part of the Ss may have been a necessary condition which determined the efficacy of Heckelman's technique. Ultimately it might also have provided speed limitations for children with normal reading capacities.

The Dunn Study

Dunn (1954) studied the qualitative differences and similarities between EMRs and normals of the same mental age on certain aspects of the reading process. Dunn offers empirical support for the special educator's anecdotal descriptions of the EMRs' reading deficit and some insights into the specific character of the problem we are trying to subvert.

Dunn's subjects were 20 mentally retarded white boys attending special classes in Danville (Illinois), and 30 mentally normal white boys who were selected at random from regular classes in elementary schools in the same school system. Stanford-Binet mental ages of both groups ranged from 8-0 to 10-0. Each subject received tests of reading, arithmetic, visual skills, and efficiency handedness, pure-tone, and home adjustment.

Dunn found that on tests of both oral and silent reading achievement, the EMRs were inferior to normals of the same MA. This finding is consonant with the results of many prior studies (Andrus, 1942; Burt, 1937; Hill, 1939; Hoyt, 1924; Kelly, 1934; Merril, 1918; Mullen, 1952; Renshaw, 1919; Scarborough, 1951; Witty and McCafferty, 1930) which have suggested that EMRs do not read up to MA expectancy.

Dunn's unique contribution, however, was his analysis of the reading error patterns which characterized the two groups:

> Faulty Vowels and Consonants. Retardates make more vowel errors than normals, but there was no difference in consonant errors. Such errors were interpreted (Monroe, 1932) in terms of inadequate auditory discrimination skills

and grasp of phoneme-grapheme correspondence relationships. Dunn offered three factors which may influence the EMR's relative inadequacy in handling vowels:

(a) Retardates tend to focus upon the initial grapheme as a clue toward recognition, and most English words start with consonants.

(b) Consonant sounds are learned first by EMRs.

(c) Vowel graphemes represent a greater number of phonemes than consonants and are thus inherently more ambiguous as a signal for a specific response.

Both groups fell "outside the normal range" with respect to faulty consonants so there may be a deficit in handling consonants as well as vowels.

Reversals. In line with prior findings (Mintz, 1946; Kirk, 1934; and Kirk and Kirk, 1935), Dunn found that EMRs did not make an inordinate number of reversal errors relative to the normal group or the test norms.

Addition of Sounds. No significant effects obtained.

Omission of Sounds. Retardates made an inordinate number of sound-omission errors. Dunn offered the following explanation:

> "...the mentally retarded tended to work through the initial sound of the word and guess at the remainder of it. Since their speaking vocabulary is limited to simpler, less complex words (Chipman, 1935), they tended to give words shorter and simpler than the printed word. Upon analysis, these shorter and simpler words yielded an excessive number of omissions of sounds. Too, they tended to skip endings which increased the tendency toward excessive omission of sounds." (p. 52)

Substitution of Words. No significant effects obtained.

Repetitions. Retardates made fewer repetition errors than the normals. Dunn cites the opinion of Monroe (1932), who contended that when the child makes errors which disrupt the meaning of the passage, he tends to re-read a portion of the text in an attempt to discover the errors. (p. 133).

Dunn suggested that EMRs make fewer repetitions because "...they were not as concerned with 'meaning' as with reading words. Thus, they tended to be 'word-by-word' readers and 'plodders'." (p. 53).

Addition of Words. EMRs made fewer addition-of-word errors than the normals. Such errors are indicative of adequate comprehension and of elaboration of that meaning in terms of personal experience. The EMR tends neither to comprehend what he reads nor to integrate the process of reading with his prior experience.

Omission of Words. No significant effects obtained.

Words Aided and Refused. EMRs made more errors of this type than normals. This finding may be attributed to the retardate's inferior word attack skills and his knowledge that he is inadequate in this respect.

Use of Context Clues. Normals were superior to the EMRs in the use of context clues on a test which would today be described as involving a "Cloze" technique. The EMR cannot use the redundancy of the language to infer an appropriate word.

Sound Blending. The two groups did not differ in the ability to blend sounds. Dunn suggested that the groups did not differ with respect to the ability to profit from instruction in phonics.

Visual Motor Functioning and Silent Reading Comprehension. The two groups did not differ with respect to the number of fixations or regressions or with respect to silent reading rate (as measured by a ophthalmograph). A test of comprehension favored the normal group. Dunn pointed out that:

> "If the retarded subjects did read the passage, which is highly problematical, they showed little understanding of what they had read. It seems probable that they tended to look at letters and words, and then move on across the line with good left to right orientation. This would explain the good eye movement record and the poor comprehension score." (p. 57)

Phrase and Word Recognition, Timed and Untimed. Tests of phrase and word recognition favored the normal group. Performance of both groups was superior under the untimed condition.

Listening Comprehension. The two groups did not differ with respect to listening comprehension, which was interpreted as an estimate of reading capacity.

School Achievement. The two groups did not differ in arithmetic computation, but the EMRs were inferior to the normals in arithmetic reasoning and oral spelling.

Auditory Acuity and Visual Efficiency. Dunn's findings suggest that there is a greater incidence of inadequate auditory acuity and visual anomalies in the EMR population.

Dunn's data support the view of the EMR as one who has learned to read but cannot read to learn. In spite of not being able to differentiate and not responding to many salient parameters of the graphemic code, the EMR does manage to transduce some information from printed to spoken form. But he doesn't get the message. The reading process is devoid of functional utility.

No attempt will be made here to "explain" these findings. This account of the EMR's reading behavior has been presented to support a set of assumptions which had a direct bearing upon the design and interpretation of the present research :

The EMR's command of (a) requisite visual motor functions, (b) many salient parameters of the graphemic and phonemic codes employed in our language, (c) phoneme-grapheme correspondence rules, and (d) oral language are such that: (1) the assumption that the EMR can profit from bimodal stimulation was tenable; (2) responses to visual stimulation would not necessarily interfere with the process of obtaining meaning from the auditory message; (3) as a function of contiguity, the process of obtaining meaning from the auditory modality would transfer to the visual stimulus input; (4) as a function of contiguity, further development in specific word recognition and word attack skills might obtain; (5) there was no a priori reason to expect that oral reading on the part of the child would be necessary in order to facilitate such transfer; (6) it might well be that imposing the task of oral reading upon the EMR may serve to induce negative motivational factors, focus his attention upon an ancillary and essentially gratuitous task parameter, and actively interfere with the primary goal of comprehending the message.

The above considerations were the primary reasons for departing from the procedures employed by Heckelman (1964). The relative efficiency of group instruction was a secondary determinant of the course of action.

The Spicker Study

Spicker (1963) studied the listening comprehension and retention (after one week) of 44 EMRs and 44 intellectually normal children of the same MA as a function of passage difficulty (3rd, 5th, and 9th grade), and rate of presentation (125, 175, 225, and 275 wpm). The material was tape recorded at 175 wpm and was electronically expanded or compressed by means of a Tempo Regulator (the device employed in the present study) to yield the other rates. Passage difficulty was treated as a within Ss source of variance.

Spicker found that:

(1) No significant difference in comprehension was detectable between the two populations;

(2) In general, retention on the part of the EMRs was superior to that of the normals;

(3) The 125 and 175 wpm rates yielded superior comprehension on the part of both groups than did the 225 and 275 wpm rates;

(4) There were no significant differences in comprehension between the 3rd and 5th grade material, while comprehension of the 9th grade passage was inferior to that which obtained with the other passage;

(5) On the basis of a non-significant trend, the 125 wpm (expanded) rate was more effective for the retardates than for normals, and the converse was true for the 175 wpm rate.

Spicker's findings suggest that retardates can comprehend auditory materials at rates of at least 175 wpm. This is well beyond their reading rate. In the case of bimodal stimulation, the efficiency gained by increasing rate may be counteracted by interference caused by a breakdown in the reading process. Travers, et al, (1964) found the audio-visual (bimodal) condition inferior to unimodal stimulation when the stimulus exposure duration time was limited to one second. As we are concerned with potential increments in word recognition and word attack skills as well as listening comprehension, the facilitation related to the effects of distributed versus massed practice found in studies of rote associative learning are deemed relevant. Such findings would suggest that slow rates of presentation would enhance the development of reading skills as well as preclude the possibility of overloading the EMR's information processing channel capacity. It was thus clear that the first order of experimental business must be to examine the effects of rate in conjunction with bimodal presentation upon the processing of information.

Experiment I

The three phases of Experiment I were designed to investigate parameters relevant to the determination of experimental procedures to be employed in the major treatment study (Experiment II). This initial work was performed with institutionalized mental retardates enrolled in the Mary Zeigler School of the District (D. C.) Children's Center, Laurel, Maryland, who met all of the following criteria:

(a) CA of 11 years or more;
(b) IQ of 45 or more;
(c) Absence of major sensory deficit;
(d) Teachers' rating of ability to profit from the treatment, and/or a recent Metropolitan Achievement grade equivalent score of 1.5 or better.

(Recent reading achievement test data were available on a substantial number of children; the unavailability of such data was not a basis for exclusion of Ss.)

The names of those selected for inclusion were submitted to the professional staff of the school in order to determine the availability of the children during the experimental period; and to exclude the children who, in the professional opinion of the Children's Center staff, could not profit from exposure to the experimental treatments, or had not been able to respond adequately to the achievement tests previously administered.

Experiment I was conducted during an interim period between the close of summer school and the start of the fall school term (August 17 through September 3, 1965).

Fifty-eight children were identified as being available and as having the potential of profiting from the experience. The 58 Ss were randomly assigned to four groups, two groups of 14 Ss and two groups of 15 Ss each. One S from one of the larger groups was lost due to disciplinary action of school personnel. Data on one S from the other larger group, randomly selected, was excluded from tabulation and analysis. Thus, each experimental group consisted of 14 Ss. Descriptive data on each group is presented in Table 1.

Each group of Ss was assigned to a fixed daily treatment period as counterbalancing for diurnal effects would have imposed inordinate demands upon the DTS staff. The treatment period time for each group was as follows:

Group I:	8:30 to 10:00 a.m.	Group III:	1:00 to 2:30 p.m.
Group II:	10:00 to 11:30 a.m.	Group IV:	2:30 to 4:00 p.m.

Prior to the commencement of Phase I, the "Speed" portion of the Pressey Diagnostic Reading Test and the "IOTA Word Discrimination" subtest of the Monroe Diagnostic Reading scales reproduced (by multilith) in 14-point IBM "Macrotype" bold-face, sans serif (speech reading) type face was administered to all Ss.

Phase 1, Experiment I

Purpose and Procedures

Phase 1 of the experimentation was designed to test the effect of rate of auditory pacing (involving the simultaneous presentation of the visual material) upon the comprehension of material which was within the general independent reading level of the Ss.

Four equated passages from Level A of the McCall-Crabbs standardized reading exercises were prepared in spirit-duplicated, 14-point IBM "Macrotype" type face for visual presentation. An additional passage from Level A of the McCall-Crabbs reading passages was selected to serve as practice material.

Table 1

Summary Data on Groups in Phase 1, Experiment I

Group		IQ	Sex M	Sex F	CA (in mos.)	Pressey	IOTA	Av. Read. Gr. Equiv.
I	$\bar{X}$	60.6			174.6	5.1	18.4	2.6
	s	7.0			18.2	4.1	8.2	.68
	N	14	10	4	14	14	14	9
II	$\bar{X}$	57.1			183.3	5.2	24.6	2.8
	s	7.2			20.9	5.9	9.7	1.17
	N	14	7	7	14	14	14	10
III	$\bar{X}$	54.8			183.8	5.8	26.3	3.0
	s	7.4			21.6	5.2	7.3	.69
	N	11	4	10	12	13	14	9
IV	$\bar{X}$	57.0			182.5	7.4	21.6	3.0
	s	5.6			21.8	6.5	11.8	.78
	N	14	6	8	14	14	14	9

After the testing described above, on the same day of the testing, the practice passage was presented to each of the four groups of Ss to insure that each S understood the nature of the task. On the next day, one of the four passages was presented after a copy of the visual material had been passed out to each S. The Auditory material was presented by means of a Tempo Regulator (a device which can vary rate without pitch distortion), and rate was varied according to the dictates of the design (see below). After presentation of the passage, the visual copy was collected and a set of 10 questions, spirit-duplicated in Macrotype, was passed out. E read each of the questions, reciting each of the four multiple choice alternatives twice. The remaining passages were presented similarly on succeeding days.

The auditory materials were recorded at approximately 50 to 60 wpm by a male Caucasion speaker who increased the length of pauses between words in order to reduce rate. This technique was employed in conjunction with the Tempo Regulator in order to reach the extremely slow rates desired for this experiment.

A Lindquist (1953) Type II "Mixed" (Latin Square) design was employed in Phase 1. Passage was counterbalanced across rate to insure that any difference in passage difficulty or practice effects due to prior amount of exposure would be controlled in overall rate comparisons. This procedure held order of passage presentation constant and thus confounded rate with degree of expansion or compression (the extent to which the machine had been employed to achieve the rate). The design is shown in Table 2.

Results

Criterion data (number of correct responses) for the four groups are presented in Table 3.

The summary table for the Latin Square analysis of variance (ANOVA) is presented as Table 4.

Table 2

The Design of the Experiment Employed in Phase 1

Group	Condition			
	Expanded 50% (28-30 wpm)	Expanded 20% (44-48 wpm)	Normal (55-60 wpm)	Compressed 20% (65-71 wpm)
I	A2(1)	A5(2)	A6(3)	A7(4)
II	A7(4)	A6(3)	A5(2)	A2(1)
III	A5(2)	A2(1)	A7(4)	A6(3)
IV	A6(3)	A7(4)	A2(1)	A5(2)

Note.--The passage (and order of presentation) are expressed as the Latin Square variable.

Table 3

Raw Score Means and Standard Deviations on Comprehension Questions by Groups and Conditions

Group		Condition				
		Exp. 50%	Exp. 20%	Nor.	Comp. 20%	Total Group
I	$\bar{X}$	6.3	5.8	6.1	7.1	25.3
	s	1.9	2.8	2.0	1.8	7.5
II	$\bar{X}$	7.0	5.9	5.6	6.2	24.7
	s	2.4	2.6	2.6	1.9	7.9
III	$\bar{X}$	6.4	5.7	7.3	5.4	24.9
	s	2.7	2.5	2.2	2.1	8.1
IV	$\bar{X}$	5.6	6.6	7.1	6.1	24.7
	s	1.7	1.7	2.0	2.9	24.9
	$\bar{X}$	6.3	6.0	6.5	6.2	

Table 4

Summary Table for Latin Square ANOVA, Phase 1, Experiment I

Source	df		MS	F
Between Subjects		55		
Groups	3		.40	-
Ss w/Groups	52		11.34	
Within Subjects		168		
A Condition	3		15.81	2.30*
B Passages	3		2.78	-
(AB)	6		12.55	1.82
Error (B)	156		6.88	
Total		223		

$*F_{.90}(3/156) = 2.12$

The analysis of variance indicated that there was no difference in comprehension at the 5% level of significance as a function of rate across the range of 20 to 60 words a minute with material within the general reading level of the Ss. A trend at the 10% level was obtained. A Duncan Multiple Range Test (Edwards, 1965) showed that the normal (55-60 wpm) presentation was superior to the 20% expanded condition. (It should be remembered, however, that rate was confounded with degree of expansion/compression.) No passage or group differences were identified.

Phase 2, Experiment I

Upon completion of the first phase of the experimentation, Ss were given the option of continuing to participate in the program. Thirty-five Ss elected to continue. This was interpreted as a gross measure of general motivation. No attempt was made to persuade Ss to continue. A comparison of descriptive data (IQ, CA, Ach, etc.) and total criterion data (correct responses on all four passages) between "motivated" and "unmotivated" Ss is presented in Table 5. There were no significant differences between the means of the two groups on any measure ("t" tests at 5% level of significance).

The retardates who expressed a willingness to continue constituted the subject population employed in Phases 2 and 3 of Experiment I. The primary purpose of Phases 2 and 3 was to supply a basis for judgment as to what procedures should be employed in Experiment II.

Procedures

The design of Phase 2 involved four factors:

Within S Comparisons

A. Passage Difficulty

a_1--Easy

a_2--Hard

Table 5

Comparison of Descriptive Data on Motivated and Unmotivated Ss in Phase 1, Experiment I

Data Source		Motivated Ss	Unmotivated Ss
IQ	$\bar{X}$	58.3	56.3
	s	7.2	6.6
	N	34	19
CA	$\bar{X}$	180.7	181.2
	s	18.3	24.5
	N	34	20
Pressey	$\bar{X}$	5.9	5.9
	s	5.2	5.8
	N	35	21
IOTA	$\bar{X}$	20.9	25.7
	s	9.6	9.3
	N	35	21
Av. Read. Grade Level	$\bar{X}$	2.8	2.8
	s	.77	.98
	N	22	15
Criterion Data	$\bar{X}$	25.6	23.7
	s	7.1	7.4
	N	35	21

B. Modality of Testing

b_1--Listening and Reading (LR)

b_2--Reading only (R)

Between S Comparisons

C. Rate of Presentation

c_1--Normal (about 100 wpm)

c_2--Expanded (50 percent slower)

D. Modality of Presentation

d_1--Listening and Reading (AV)

d_2--Listening only (A)

The order of within S experimental conditions was as follows:

Day	Difficulty (A)	Test Mode (B)
1	Easy	Listening and Reading
2	Hard	Listening and Reading
3	Easy	Reading Only
4	Hard	Reading Only

The content of the four passages involved space and earth science. The materials were adapted from sources designed for use at a junior high school level. Reduction of vocabulary load was the primary means of manipulating the readability of the difficult passages. Reducing the difficulty of the hard passages while maintaining fidelity of content entailed drastic reduction of sentence length and complexity. Writing style was sacrificed to this end.

The sentences were short but choppy. Sentence structure was almost invariant: (modifier)-noun-verb-(modifier)-noun. Of necessity, there was heavy reliance upon the repetition of substantives and the use of indefinite pronouns.

The number of questions (four-alternative, multiple-choice) used to test comprehension of the passages employed in Phases 2 and 3 varied from 15 to 40. The percent of correct responses by each S served as the unit of analysis. All questions were developed by a skilled, experienced specialist in item construction.

Due to the unequal number of Ss in each group (of Phase 1) who were willing to continue in the program, it was necessary to redistribute Ss from two of the groups. Ss were randomly selected from the two larger groups and randomly assigned to the two groups with low remaining numbers of Ss. Of the 35 Ss who expressed a willingness to continue, one was lost due to school disciplinary action and another was hospitalized. Data on one other S (randomly selected) was excluded from tabulation and analysis. Thus, data on eight Ss in each group (N=32) were tabulated and analyzed. Descriptive data on these Ss are presented in Table 6.

The procedures employed in Phase 2 were essentially the same as those used in Phase 1 with the exception that no visual material was presented to those Ss in the "Listening Only" mode of presentation condition (d_2) and that E did not read the questions under the "Reading Only" (b_2) mode of testing. Descriptive data on passages and rates of presentation employed in Phase 2 are presented in Table 7.

Results

Summary data on the percent of correct responses is presented in Table 8. These data were treated in a four-dimensional analysis of variance. The summary table is presented in Table 9.

This ANOVA is of most value in offering some insight into the potential influence of the variables of interest, as the strictly statistical interpretations are somewhat obscure. To aid in interpretation, the significant interactions are plotted in Figures 1, 2, and 3.

Table 6

Descriptive Data on Ss in Each Group Involved

in Phase 2 of Experiment I

Data Source		Group and Condition (Between S Comparisons)			
		I	II	III	IV
		AV(c_1d_1) Normal	A(c_2d_2) Expanded [a]	A(c_1d_2) Normal	AV(c_2d_1) Expanded
IQ	$\bar{X}$	62.0	57.1	55.3	58.8
	s	7.7	6.7	9.3	4.2
	N	6	7	7	8
CA	$\bar{X}$	174.5	179.1	183.1	186.9
	s	20.6	14.2	20.4	19.7
	N	8	7	7	8
Pressey	$\bar{X}$	3.8	5.3	6.9	8.4
	s	2.5	3.4	6.5	7.2
	N	8	7	8	8
Monroe	$\bar{X}$	16.6	22.9	22.5	22.0
	s	6.5	10.7	8.1	13.6
	N	8	7	8	8
Reading Achievement	$\bar{X}$	2.4	2.9	3.0	3.2
	s	0.73	0.41	0.75	1.0
	N	4	5	6	5
Total Correct Responses (Phase I)	$\bar{X}$	24.4	30.3	26.0	23.2
	s	5.5	4.4	7.7	6.4
	N	8	7	8	8

[a]The c_2d_2 cell includes one S lost to Phase 1 due to disciplinary action by school personnel.

Table 7

Passage Grade Level and Rates of Presentation

Employed in Phase 2

Condition			Grade Level[a]	Approximate Presentation Rate (wpm)
Difficulty	Test Mode	Presentation Rate		
Easy	LR	Nor (a_1, d_1, c_1)	3.2	108
Easy	LR	Exp (a_1, d_1, c_2)	3.2	44
Hard	LR	Nor (a_2, d_1, c_1)	6.1	101
Hard	LR	Exp (a_2, d_1, c_2)	6.1	55
Easy	R	Nor (a_1, d_2, c_1)	3.3	96
Easy	R	Exp (a_1, d_2, c_2)	3.3	43
Hard	R	Nor (a_2, d_2, c_1)	5.7	107
Hard	R	Exp (a_2, c_2, d_2)	5.7	49

[a]Average of four samples according to Dale-Chall or Spache readability formulae.

Table 8

Summary of Criterion Data (Mean Percent Correct Responses)

from Phase 2 of Experiment I

Rate (c) and Mode (d) of Presentation		Modality of Testing: Listening and Reading (b_1) Easy (a_1)	Listening and Reading (b_1) Hard (a_2)	Reading Only (b_2) Easy (a_1)	Reading Only (b_2) Hard (a_2)
Group I (N = 8)					
Normal (c_1)	$\bar{X}$	39.0	27.9	26.2	23.8
AV (d_1)	s	3.7	7.1	9.2	6.9
Group II (N = 8)					
Expanded (c_2)	$\bar{X}$	42.0	32.5	31.9	27.8
A (d_2)	s	14.0	15.6	9.6	5.1
Group III (N = 8)					
Normal (c_1)	$\bar{X}$	34.0	25.0	25.6	30.9
A (d_2)	s	20.2	16.1	15.2	12.0
Group IV (N = 8)					
Expanded (c_2)	$\bar{X}$	37.5	27.1	45.0	28.2
AV (d_1)	s	10.7	11.9	12.8	7.0

Table 9

Summary Table for Four-Dimensional ANOVA

Employed in Phase 2 of Experiment I

Source	df		Ms	F
Between Ss		31		
C Presentation Rate	1		834.87	2.26
D Presentation Modality	1		6.16	--
CD	1		6.48	--
Ss w/Groups	28		369.24	
Within Ss		96		
A Difficulty Level	1		1,604.75	36.33**
B Test Modality	1		286.26	2.61
AB	1		211.61	3.54*
AC	1		315.95	7.16**
BC	1		66.61	--
AD	1		241.78	5.47**
BD	1		55.18	--
ACD	1		14.58	--
BCD	1		657.03	7.09**
ABC	1		252.17	4.22**
ABD	1		180.41	3.02*
ABCD	1		11.28	--
A x Ss w/Groups	28		44.16	
B x Ss w/Groups	28		92.61	
AB x Ss w/Groups	28		59.69	
Total		127		

$*F_{.90}\ (1, 28) = 2.89$

$**F_{.95}\ (1, 28) = 4.20$

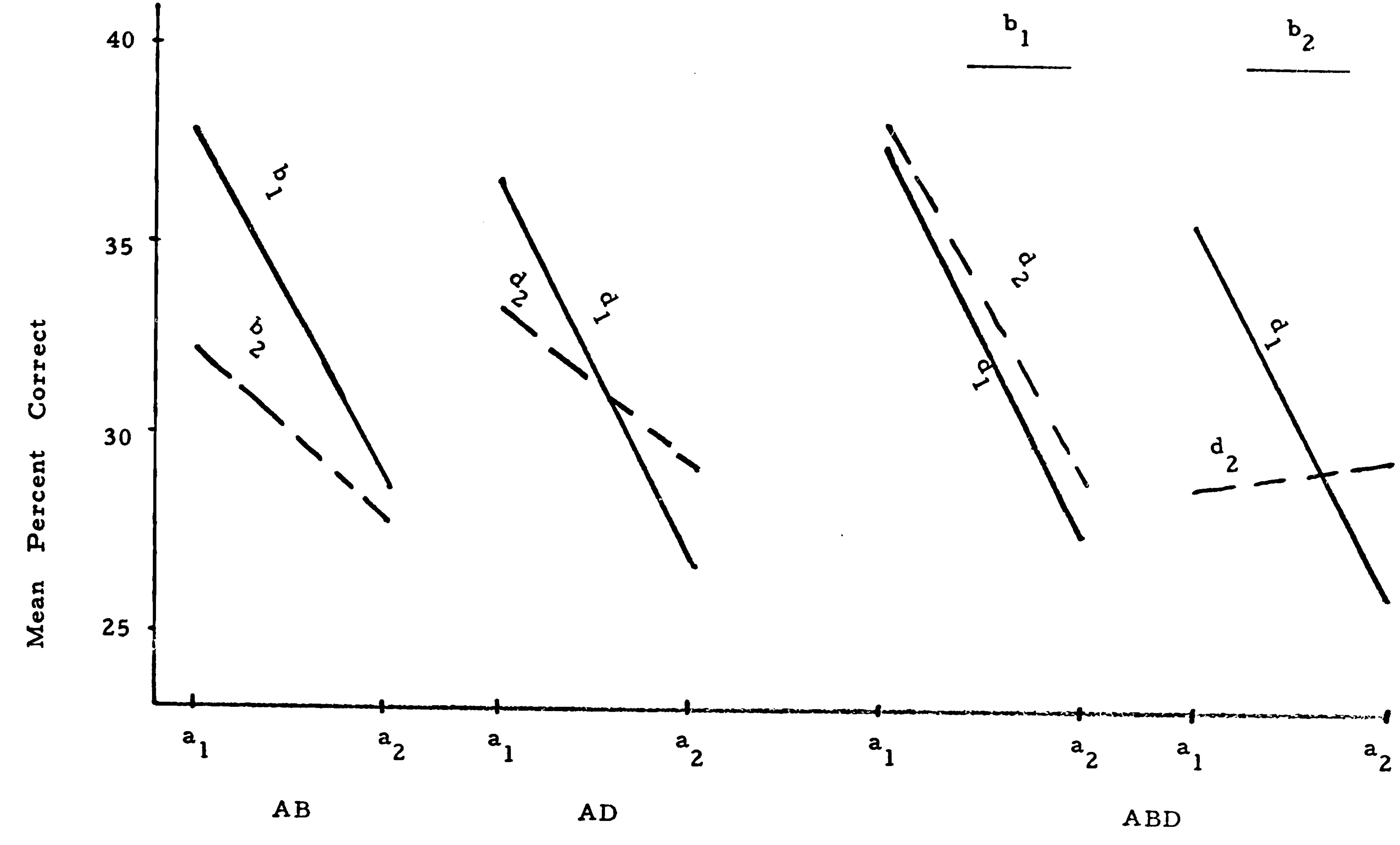


Fig. 1 First and second order interactions among difficulty (A), presentation mode (D), and testing mode (B).

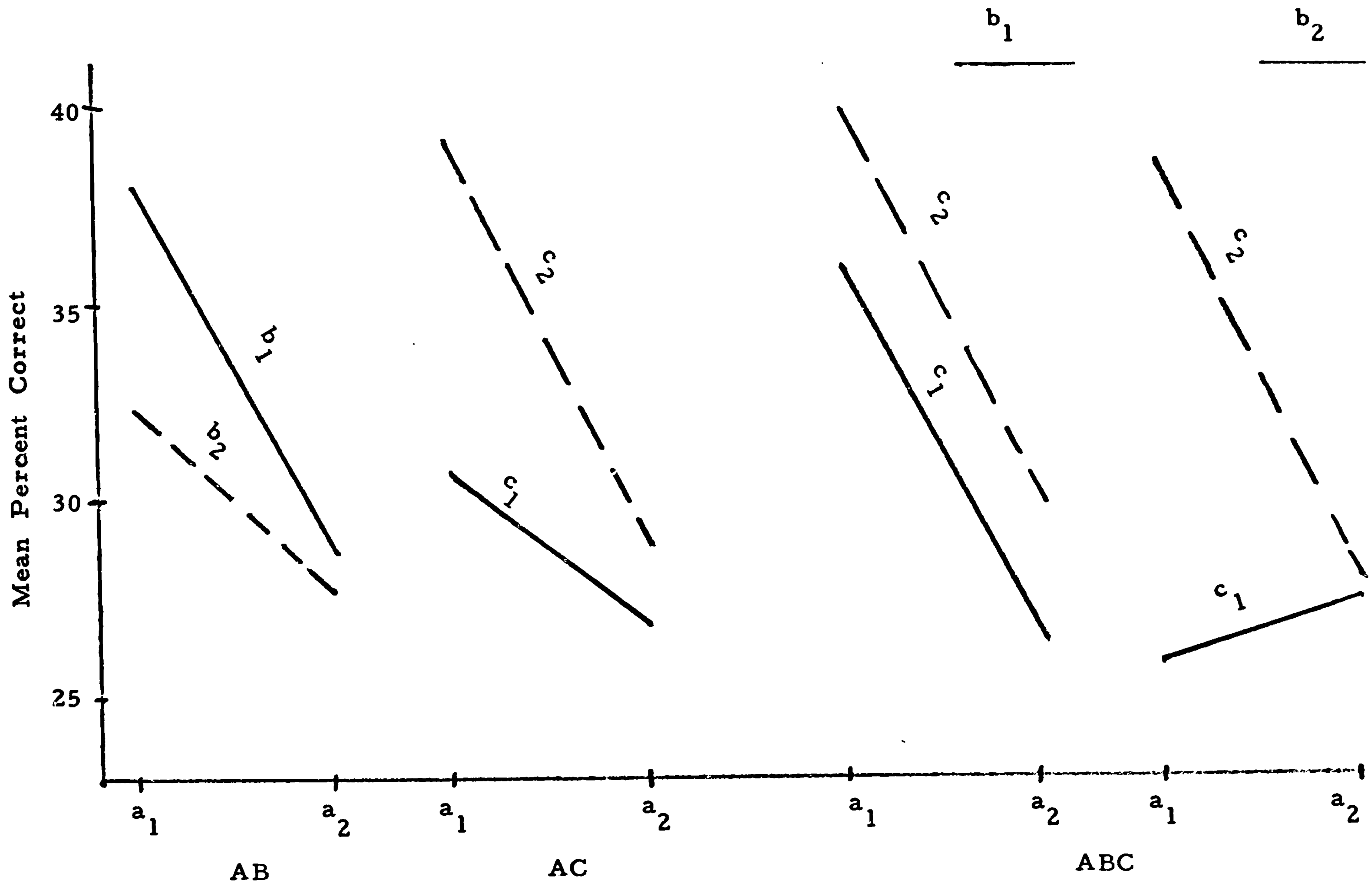


Fig. 2 First and second order interactions among difficulty (A), presentation rate (C), and testing mode (B).

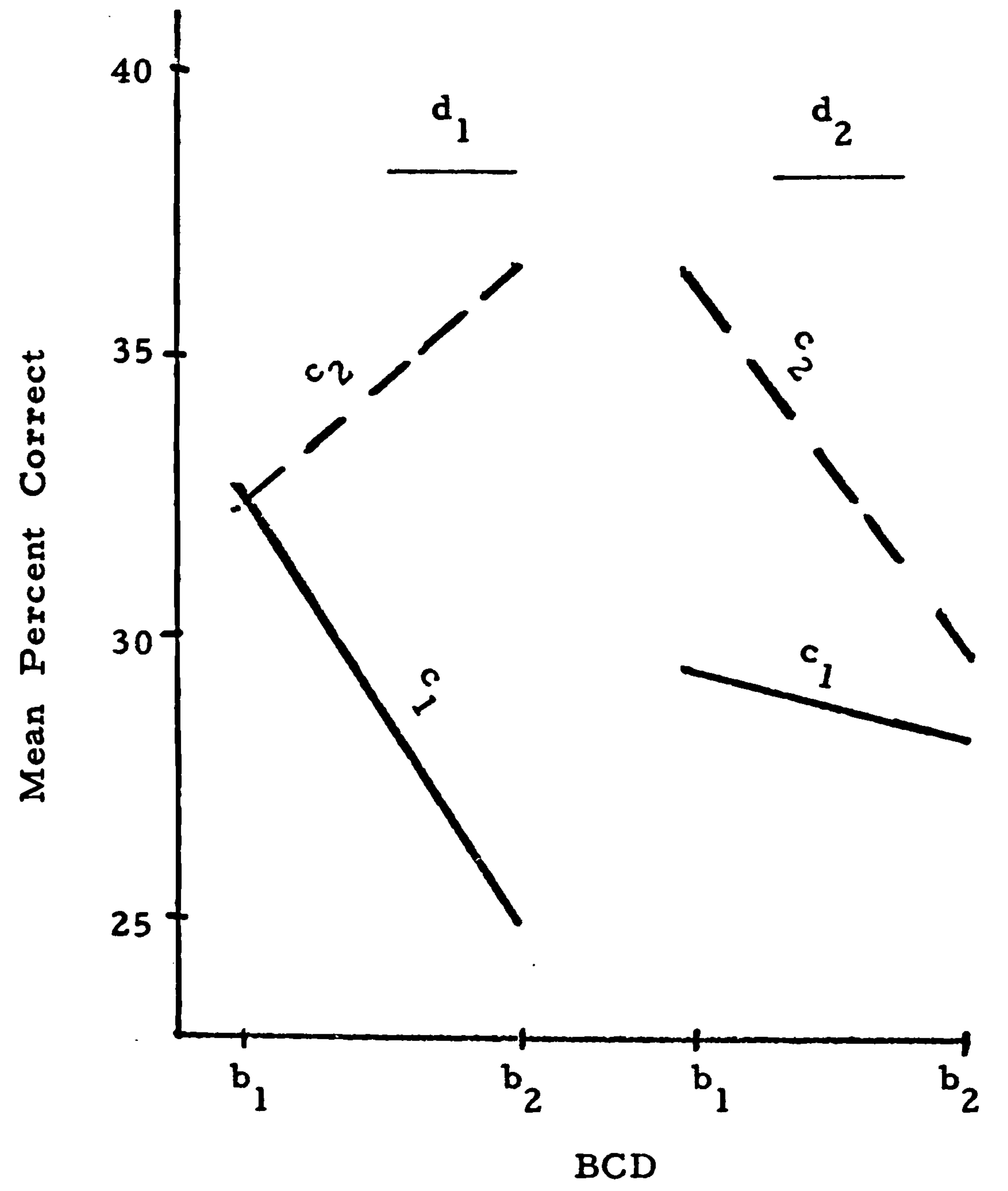


Fig. 3 Second order interaction among testing mode (B), presentation rate (C), and presentation mode (D). (First order interactions BC, BD, and CD were not significant.)

Figures 1 and 2 show that, as might be expected, comprehension was generally better on the easier passages. Both audio-visual techniques (testing and presentation) appeared to work relatively better with easy material than hard material, and expanded presentation rate was also comparatively more effective with easy than with hard material. These findings might have been a function of the fact that performance on the hard rate was barely above a chance level for all cells. The Listening Only presentation combined with the Reading Only test seemed to have been particularly low for the easy material (slightly lower than for the hard material). This finding perhaps suggests an initial confusion since the Listening Only group received first easy material then hard material on the next day under the Reading Only test condition.

Figures 2 and 3 suggest a general superiority of the expanded rate which was about half the speed of the normal presentation rate. Figure 2 shows that the normal rate was particularly poor for easy material under the condition of Reading Only testing. Figure 3 shows that the expanded rate of presentation was especially effective where the AV presentation mode and the Reading Only testing mode were used.

In summing up these findings, it seems clear that comprehension scores tended to be higher under conditions of easy material, of expanded presentation rate, and of bimodal testing (with exceptions as noted above in discussing the interactions). Various interactive cancellation effects preclude a general statement about presentation mode.

Phase 3, Experiment I

The purpose of Phase 3 was to examine the effect of rate of presentation independent of the degree of expansion or compression. The same groups that participated in Phase 2 took part in this phase. The design involved three factors:

Within S Comparison:

C Rate of Presentation (slow, medium, fast)

Between S Comparison:

A Presentation Modality (Audio versus Audio Visual)

B Recording Rate (approximately 55 versus approximately 110 wpm)

The experimental design is shown in Table 10.

Table 10

Experimental Design Employed in Phase 3, Experiment 1

Recording Rate	Presentation Rate		
	Slow (56-65 wpm)	Medium (65-75 wpm)	Fast (95-105 wpm)
	Passage I	Passage II	Passage III
Slow	Normal	25% Compressed	50% Compressed
Fast	50% Expanded	25% Expanded	Normal

Note.--Cell entries indicate degree of machine manipulation needed to equate rates of presentation. This design was replicated across the presentation mode (AV vs. A) dimension.

An attempt was made to select three passages of approximately equal length, recording rate and estimated readability. Due to a confusion between two passages with the same title, this condition was not met. Passage III turned out to be easier than the other two, and the initial recording rate involved in the "slow" ordering of this passage was well below the target value of 55 wpm.

As this error confounded the results with passage difficulty, the results of the ANOVA are not presented here. Suffice it to say that the results showed that performance was superior on the fast rate (easy passage) and that they may well be attributed to the differential difficulty of the passages.

Discussion of Experiment I

The general purpose of Experiment I was to supply information which might assist in the determination of procedures to be employed in Experiment II. Of primary interest were the effects of:

(1) Rate of presentation

(2) Rate of recording

(3) Bimodal (AV) versus unimodal (A) stimulation

(4) The distortion inherent in the expansion-compression process on the EMR's comprehension of the passage as reflected by multiple choice tests over the content of the material. Implications drawn from each phase of Experiment I are discussed.

Phase 1, Experiment I

The results of Phase 1 suggested that there was no difference in comprehension of the content of bimodally (AV) presented material within the general reading level of the Ss across the range of 20 to 60 wpm. There was some evidence of a slight restriction in the range of the criterion data, but this restriction did not seem to be responsible for the absence of a rate effect since mean comprehension did not vary as a monotonic (positively or negatively) function of rate. It was concluded that any distortion inherent in the expansion-compression process did not have a significant effect on comprehension.

Phase 2, Experiment I

The implications of the Phase 2 results were quite influential in the determination of subsequent experimental procedures:

Passage Difficulty. The EMR's comprehension appeared to be enhanced by re-writing intrinsically difficult material in a dull, choppy, repetitive style which most "normal" children (and teachers) would find quite intolerable. This is in sharp contrast to the usual procedures employed in developing "high interest, low vocabulary" material for this population, wherein a reduction in vocabulary load is the major technique employed in making the material more readable. A sharp restriction in sentence length and complexity (as in the present study) is usually sacrificed in order to maintain an "interesting style." It may be, however, that the greater syntactic redundancy of the dull, repetitive style may be more facilitative of the EMR's comprehension than attempts to enhance the motivational properties of the passages. Nevertheless, this lead was not followed up in Experiment II since there was greater concern over establishing some basis for generality across the stylistic and content factors which characterize material in current use and a desire to insure that the effects of modality and rate of presentation would not be limited to material written in an extremely esoteric style of writing.

Test Mode. The negative reaction of the Ss to the reading only tests of comprehension after prior use of both channels to present the test stimuli raises the issue of the validity of treating test mode (unimodal versus bimodal) as a within S source of variance. It was decided to omit reading only tests of comprehension in order to preclude the possibility of negative motivational effects being engendered in this fashion. Such effects would tend to invalidate any comparison of the effects of reading versus listening comprehension and might well extend into subsequent teaching periods.

Presentation Rate. The general superiority of comprehension associated with expanded (slow) presentation of the auditory material (Figure 2) obtained in spite of the fact that "normal" presentation was at a rate quite consonant with that of oral reading on the part of teachers. This finding added some empirical support to a theoretical position, cast in terms of the effects of massed versus distributed effects in rote associative learning and the slow reading rates and assimilation of the EMR, that slow presentation might facilitate comprehension. It was decided to use slowed speech in Experiment II.

This decision, however, created another problem: that of the somewhat less acceptable vocal quality of the output under machine expansion. While there is no pitch distortion, and the speech is intelligible, it was feared that prolonged exposure to the expanded speech might have an adverse motivational effect. This problem was approached by reducing initial recording rate by having the speaker increase the pauses between words and clip the length of vowel sounds in a rather staccato delivery technique. The Tempo Regulator, under expansion, compensates for the recording technique and further increases the length of pauses between words. By this means it was possible to reduce the rate of vocal presentation as low as 20 to 25 wpm without loss of relative clausing, intonation or inflection patterns, or the induction of poor vocal quality.

Thus, these tapes were "designed" for expansion. As a consequence, there was a sharp reduction in the potential rate range of such tapes obtainable by use of the Tempo Regulator relative to tapes wherein a higher initial recording rate is employed. In addition, the special tapes were not as satisfactory under compression as much of the redundancy associated with the longer temporal duration of vowels had already been eliminated. We were faced with another potential adverse motivational factor: inadequate flexibility in rate range to match the needs and preferences of the Ss. It was decided to record the passages at a higher rate as well, employing a standard vocal delivery, in order to be prepared for this potential problem.

Presentation Mode. The effects of unimodal versus bimodal stimulation on comprehension were not clear cut with respect to the existence of an effect nor of the direction and degree of such an effect. In light of this, it was decided not to require Ss exposed to bimodal stimulation in Experiment II to attend to both channels. It was assumed that:

(1) At present, the EMR is the best judge of the most efficacious channel or channel combination in a particular situation;

(2) If motivated, the child would attend to the most profitable sources of information;

(3) If unmotivated, the child would not respond positively to attempts to coerce bimodal attention.

Thus, it was concluded that permitting differential attention to bimodal stimulation would not be apt to have a deleterious effect on comprehension. The study is somewhat criterion bound, however, in that a similar rationale will not hold with respect to increasing reading skills. If there is to be an increase in such skills as a function of bimodal stimulation, attention to both channels is essential.

Phase 3, Experiment I

The results of Phase 3, in line with the prior phases, offered little evidence that the mode of presentation (unimodal versus bimodal) was a potent source of variance in the determination of the EMR's comprehension. This observation must be qualified by pointing out that a visual only exposure condition was not employed as a control. There is ample evidence in the literature to support the position that the EMR's reading comprehension is vastly inferior to his listening comprehension.

It might be well, at this point, to present the rationale for the omission of a visual only exposure condition in Experiment II:

(1) To test the effects of stimulation modality on comprehension, it is necessary to hold the content of the passage and the test constant across mode of exposure. The marked disparity between the EMR's reading and listening comprehension is so extreme that passages appropriate in difficulty for the reading group would be too easy for the listening group, and material appropriate for the listening group would be too difficult for those reading such material.

(2) Thus, performance on the criterion tasks would be primarily a function of artifacts associated with difficulty effects.

(3) In general, with an EMR population, the utility of a unimodal visual exposure condition would be limited to service as a baseline control.

(4) It would not be desirable to employ such a baseline control group in a case such as Experiment II wherein some 40 to 50 hours of instructional time would be devoted to pre-

senting material, the content of which was either: so easy as to be of little instructional value to any group (A, V, or AV); or so difficult as to be incomprehensible to those in the visual exposure condition.

In summary, the results of Experiment I were influential in the determination of three major decisions with respect to the establishment of the procedures employed in Experiment II:

(1) Most of the passages presented auditorially by the Tempo Regulator were presented under expansion and at very slow rates.

(2) All tests of content comprehension involved bimodal presentation of test materials to all groups.

(3) No attempts were made to coerce or cajole Ss exposed to bimodal stimulation to attend to both channels if they were not inclined to do so.

Experiment II

Purpose and Procedures

The purpose of Experiment II was to assess the effects of bimodal (auditory and visual) stimulation with redundant material upon the reading and listening skills of Educable Mentally Retardates (EMRs).

Two bimodal stimulation groups were established: one for which auditory materials were presented by a teacher while the students read the same material, and one for which the auditory material was presented by a Tempo Regulator (a device which varies the rate of presentation of taped material without pitch distortion) while the students read it. The comprehension scores over these materials were compared with those of a group exposed to auditory presentation only (by means of the Tempo Regulator).

Data collected on the two bimodal stimulation groups, based upon the pre and post-test administration of several standardized tests, were compared with those of the Ss who were exposed to the audio materials alone, and also with a group who did not participate in the experimental instructional program.

Subject selection. Ss were selected from the school population of the Rock Terrace High School, Montgomery County (Maryland) Public Schools. This school was made available through the kind cooperation of County officials and the school staff. Rock Terrace serves a population of Special Education students (principally EMRs) who, for various physical and/or psychological reasons, are unable to profit from the secondary level Special Education programs which are included in the regular high schools. The enrollment of Rock Terrace is approximately 130 students.

Pre-test data collection. The following instruments were administered to the entire school population:

(1) The Bond, Clymer, Hoyt (BCH) Silent Reading Diagnostic Tests (Form D-A) which yielded the following scores (used for analysis purposes):

(a) Recognition Pattern (BCH-RP)
Total correct on tests of recognition of words in context and in isolation;

(b) Visual Analysis (BCH-VA)
Total correct on tests of syllabication and location of word elements and roots;

(c) Phonetic Knowledge (BCH-VA)
Total correct on tests of phoneme-grapheme correspondence;

(d) Word Synthesis (BCH-VA)
Total correct on a test of the comprehension of material wherein key words are hyphenated and separated;

(2) The Durrell-Sullivan Reading Capacity Test (DSRC) (Intermediate level: Form A) which yields three scores:

(a) Total Score (DSRC-TS),
(b) Word Meaning (DSRC-WM),
(c) Paragraph Meaning (DSRC-PM);

(3) The Diagnostic Reading Tests (DRT), Survey section (K-4), Form B, Levels II and/or III as appropriate. These tests yield three scores:

(a) Total Score (DRT-TS),
(b) Word Attack (DRT-WA),
(c) Comprehension (DRT-C);

(4) The Primary Mental Abilities Test (PMA), grades 4-6, which yields six scores:

(a) Total Score (PMA-TS),
(b) Reasoning (PMA-R),
(c) Perceptual Speed (PMA-PS),
(d) Verbal Meaning (PMA-VM),
(e) Number Facility (PMA-NF),
(f) Spatial Relations (PMA-SR).

Identification of potential Ss. On the basis of data from cumulative records and the administration of the above tests, 68 Ss who met all of the following criteria were identified:

(1) SB, WISC, or WAIS IQ between 54 and 86,
(2) Chronological Age (CA) between 12 years, 11 months and 17 years, 11 months,
(3) Absense of major visual defect,
(4) Absence of known degenerative neurological disease,
(5) A grade level equivalent of at least 2.0 on the BCH-RP.

Further pre-test data collection. The following instruments were administered to the 68 EMRs identified as potential Ss:

(1) The Durrell-Sullivan Reading Achievement Test (DSRA) (Intermediate level: Form B), which yields three scores:

(a) Total Score (DSRA-TA),
(b) Word Meaning (DSRA-WM),
(c) Paragraph Meaning (DSRA-PM);

(2) The Peabody Picture Vocabulary Test (PPVT), Form B;

(3) The "speed" portion of the Pressey Diagnostic Reading Tests (PDR) which was reproduced (by multilith) in the 14-point IBM "Macrotype" bold face, sans serif

(speech reading) type face which was used in presenting the visual material during the instructional phases of the experiment. This test consists of reading material (Grade level estimates of readability, Spache Formula: 2.9; 3.6; 3.5; 3.2) in which are embedded (at intervals of approximately 50 words) binary choice comprehension items. Performance on the test is a function of both reading rate and comprehension;

(4) The IOTA and Discrimination Subtest of the Monroe Diagnostic Reading Scales, reproduced in "Macrotype" and administered as a group test;

(5) A modification of an auditory-visual integration index developed by Birch and Belmont.

Audiometric Screening. A pure-tone audiometric screening test was performed on the population of potential Ss with a Zenith model Z-200 Audiometer. Each ear was tested monaurally at 10db above threshold at the following frequencies: 4,000; 3,000; 2,000; 1,000; 750; 500; and 250 cps. In a few cases where there was a wide disparity between performance of the two ears, the series of frequencies was tested binaurally. On the basis of this screening, seven of the 68 Ss were eliminated.

Description of Experimental Groups. Four groups were constituted:

(1) A Machine-Audio-Visual (MAV) group which was exposed to bimodal stimulation of the instructional materials. The auditory materials were tape recorded and presented by means of a Tempo Regulator, a device which can vary the rate of presentation of recorded materials without pitch distortion by sampling out or sampling in minute portions of the speech record;

(2) A Teacher-Audio-Visual (TAV) group which was exposed to bimodal stimulation of the instructional materials. In this case, the auditory materials were read to the Ss by the teacher at normal speeds;

(3) A Machine-Audio (MAud.) group which was not exposed to the instructional materials through the visual modality. The auditory instructional materials were presented by means of the Tempo Regulator;

(4) A Classroom Control (CC) group which was not exposed to the experimental instructional program.

The lack of a sufficient number of Ss precluded the establishment of:

(1) A Teacher-Audio condition wherein the teachers read the material and the Ss listened without being exposed to the visual material. This would have permitted the study of unimodal versus bimodal stimulation and teacher versus machine presentation in a two-dimensional, factorial design;

(2) A Hawthorne control group taught irrelevant skills by the experimental teachers.

Constitution of Experimental Groups. The 61 potential Ss were organized according to home classroom assignment. Ss from each class were then randomly assigned to one of the four conditions (MAV, TAV, MAud, CC) with the restriction that the N of the three experimental instruction groups was 15 and that of the control group was 16.

Elicitation of Parental Consent. The parents of the 45 Ss assigned to the experimental instructional groups were contacted in order to obtain permission for including their children in the experimental program.

The parents of one child refused. One S was randomly selected from the control group and assigned to fill the vacancy. On the basis of this high parental permission rate, it is assumed that selection bias among the four groups was minimal.

General Procedures: Instructional Phases

Assignment to Teaching Period. To minimize schedule conflicts, two teaching periods, taught by two different, experienced and qualified Special Education teachers, were established for each experimental instruction condition (6 teaching periods in all). Assignment to teaching period was suited to the convenience of the children and the teaching and administrative staff of the school. As such assignment was not made on a random basis, this classification is ignored in the analysis of data, however no systematic bias seemed likely in this arrangement. Ss were assigned to a teaching period during which they would otherwise be assigned to their homeroom teacher. No S was assigned to a teaching period at such a time that they would be precluded from participating in "special" subjects such as gym, shop, music, art, work experience, etc. The N in each testing group was seven or eight.

Selection of Instructional Materials. The content of the materials employed for instructional purposes was selected and/or adopted from a wide variety of sources so as to be appropriate (in terms of interest and difficulty) for auditory presentation to EMRs. Materials for about 100 lessons were developed so that the teachers would have some latitude in selecting materials most suitable for the specific needs and interests of the experimental sample. The major sources of selection were readers, social studies and science tests employed in the intermediate grades, and special high interest, low vocabulary level materials often used in special classes for the EMR.

Development of Lesson Plans. To insure comparability in instructional goals and techniques, the two teachers jointly developed lesson plan outlines and appropriate visual materials for introducing the passage to the children. Considerable effort was expended in insuring that the format and control of these materials were such that both teachers felt comfortable in following the sequence and content outlined.

Format of Lesson Outline. The format for the preparation of the children for the presentation of the passage was as follows:

(1) The title and some broad questions which defined the instructional goals of the content of the material were presented;

(2) A series of words identified as apt to be unfamiliar were introduced for recognition purposes;

(3) A number of the "recognition" words had been previously identified as requiring definition. All such words were defined. The teacher had the option of defining a given word or eliciting a definition from the children and of defining any other word from the list of recognition words that seemed to present problems. The format of the lesson material was not so rigid as to preclude spontaneous interaction and discussion on the part of the students or teachers.

A sample lesson is given as Appendix A.

Preparation of Comprehension Questions. In some cases four-alternative, multiple-choice, comprehension questions were available from commercial sources to assess comprehension of the context of the daily lessons. In most cases, appropriate questions had to be developed. Ten questions per passage were selected from a pool of about 20 to 80 questions developed for each passage by a specialist in item construction.

Preparation of Instructional and Test Materials for Visual Presentation. The lesson outline, the comprehension questions, and the passage were typed in a 14-point, bold-face IBM "Macrotype," (speech reading) sans serif, type face which is quite similar to that of a "primary" type-writer. The copy was reproduced by an offset duplicating (multilith) process.

Preparation of Instructional Materials for Auditory Presentation. Each selection was recorded at 15 ips on a Model 728 Magnecord "Professional" Tape recorder by a male speaker who shortened the length of vowel phonemes and increased the spacing between words as this procedure seemed to enhance the vocal quality of the recording under conditions of expansion (slowing of rate below that of the original recording). Relative clausing and normal intonation and inflection patterns were maintained in almost all cases. The materials were recorded at approximately 50 and/or 100 words per minute.

Phasing of Experimental Period. The experimental period lasted for 40 school days. Each group attended one class period (50 minutes) per day. The experimental period can be defined into three distinct phases:

Phase I (Administrative) — The first teaching period was spent in explaining the experimental procedure to the children.

Phase II (Instructional) — This phase consisted of those days upon which organized lessons were presented (33 days). Phase II can be further divided as follows:

(a) Those days upon which the MAV and MAud groups heard tapes presented under expanded conditions (rates slower than that of initial recording) (25 days);

(b) Those days upon which the MAV and MAud groups heard tapes presented under compressed (rate faster than that of initial recording) or normal (rate as recorded) conditions (6 days);

(c) Those days upon which teaching was conducted but a loss of Ss due to a field trip introduced the possibility of bias into the determination of the mean of daily comprehension measures for the experimental group (2 days). Data from these days were excluded from analysis.

Phase III (Retention Study) This phase consisted of six days devoted to special procedures designed to permit the study of retention of material. In addition, on three other days involved in the instructional phase, a portion of the teaching period was devoted to gathering retention data.

Daily Experimental Procedure Employed in the Instructional Phase. A description of the sequence of events occurring daily during the instructional phase follows:

(1) The teacher introduced the lesson, following the format of the lesson outline. The TAV and MAV groups were exposed to this material in printed form as well as hearing it from the teacher. Ss recited the list of recognition words after hearing the teacher present them. The MAud group did not see this list of words. The length of time for the introduction of the passage was recorded;

(2) The passage was presented through the modalities dictated by the design (Machine vs. Teacher, Audio-Visual vs. Audio). The length of time needed for presentation was recorded;

(3) The Ss rated (on 5 point scales) each passage as to its interest value and difficulty level;

(4) The comprehension questions were presented bimodally to all groups. Printed copies of the questions were passed out to the MAud as well as the MAV and TAV groups. The teacher read each question item once and each response alternative twice. The length of time needed to present the questions was recorded;

(5) The Ss rated how well they thought they did on the questions on a 5 point scale;

(6) After collecting the answer papers, the teacher presented the questions auditorily, answered or elicited answers to them, and led a discussion as to why incorrect alternatives were inferior to correct alternatives;

(7) On the rather infrequent occasions when time remained, the teacher stimulated a discussion of the initial broad questions presented in the lesson outline.

Comparison across the Experimental Groups. As can be seen from the procedures described above, the experimental groups had many experiences in common during the instructional, as well as other, phases of the experiment. Factors common to the instructional phase:

(1) All experimental conditions (MAV, TAV and MAud)

 (a) Instructional content,

 (b) Bimodal presentation of comprehension questions (read by the teacher);

(2) MAV and TAV. Bimodal exposure to lesson outline and passage. Note: Ss were not coerced into attending to both auditory and visual materials if they did not wish to do so;

(3) MAV and MAud. Passage presentation at the same rate by mechanical means.

Presentation of Auditory Materials. The teacher read the passage to the TAV group at whatever rate was found to be comfortable for the individual teacher. In general, the rate of presentation ranged from 75 to 115 words per minute. The speech of neither teacher contained marked regional dialect features.

The auditory materials presented to the MAV and MAud groups were presented on a console which consisted of a MLR 38/15 Tempo Regulator which eliminates or repeats short (40 msec) elements of the taped material. The tape reels are driven by a Viking Model 230 Industrial Tape Transport. The auditory signal was amplified by means of a Viking PA 94 B Amplifier. Volume was adjusted daily on a subjective basis to suit varying ambient noise levels. The amplifier delivered the signal to two Electro-Voice E-V Two loudspeakers placed on either side of the classroom.

In general, the rate of presentation of auditory materials by the Tempo Regulator ranged from 40 to 65 wpm. Thus, presentation by machine was about one-half that of the teacher rate (particularly under conditions of expansion). Expansion and compression are expressed in percentage terms. Fifty percent expansion reduces the rate in half (material record at 100 wpm would be presented at 50 wpm).

Experimental Procedures Employed in the Retention Study Phase. Supplying knowledge of results to the children on the daily comprehension questions precluded using materials from the instructional phase in evaluating retention.

Materials adapted and developed by Spicker (1963) were employed in this phase of the experiment. One of three passages about animals (Grade levels: 3, 5, and 9) were presented on three successive days (the 2nd through the 4th days of the experimental period) and 20 four-alternative, multiple-choice questions (per passage) were administered. The procedures employed during the initial learning stage were essentially those employed in the instructional phase, with the exception that no introductory lesson material was presented and no discussion of the questions or the passage ensued after the data were collected. The MAV and MAud groups heard the materials under 20% expansion conditions. The 20 comprehension questions were then split on an odd-even basis. After an interval of about one month, the odd numbered questions were readministered. The ten questions (per passage) were presented on one of three successive days. About five weeks later, on the last three days of the experimental period, the even numbered questions were readministered. Immediately after the collection of this retention measure, the passages were again presented to the groups and all 20 questions were administered again to supply a re-learn index.

Post-Test Data Collection Procedures. Immediately after the close of the experimental period post-test administration of the following instruments to the four groups began:

(1) The Bond-Clymer-Hoyt (Form D-A),

(2) The Durrell-Sullivan Reading Capacity Test (Form A),

(3) The Durrell-Sullivan Reading Achievement Test (Form B),

(4) The Diagnostic Reading Tests (Form A), both levels (II and III),

(5) The Primary Mental Abilities Test,

(6) The "Speed" Portion of the Pressey Diagnostic Reading Test,

(7) The IOTA Word Discrimination Test.

Subsequent to the administration of the above tests, the questions employed in the retention study phase were administered to the control group without benefit of passage presentation to provide an empirically determined base line against which the performance of the experimental groups on the initial presentation of questions could be assessed.

Finally, near the end of the study, a Bausch and Lomb "orthorater" was employed to test all available children in the school population on:

(1) Near and far point:

(a) visual acuity,

(b) lateral phoria;

(2) Stereopsis.

A second visual test was performed within one week on a sample of the school population to provide a measure of test-retest reliability.

Results

Part I: Results of Pre- and Post-Experimental Administration of Standardized Tests

The data on Ss available for pre- and post-testing were first analyzed by employing a simple randomized ANOVA on the pre-test, the post-test, and gain score data. The Hartley Fmax ratio (Walker and Lev, 1953) was used to test for homogeneity of variance. Summary data and obtained F and Fmax ratio values are presented in Appendix B.

In no case was the F ratio significant in the analysis of pre-post criterion data. There were no significant differences among any of the groups on any of the pre-post criterion measures prior to, or as a function of, the experimental conditions. Pre-post gains for all groups were indicated in some cases, however. These data suffer from the usual limitations which obtain when one deals with EMRs using:

(1) short-term instruction,
(2) a small sample,
(3) gain score analyses,
(4) instruments designed for use in the regular grades where the recognition of correct alternatives define success.

The above limitations represent possible sources of both alpha and beta error. The lack of significant F ratios obviously precludes the possibility of alpha error. Of more pressing concern is the possibility of beta (Type II) error. A number of factors related to a lack of power were beyond the control of the investigators. However, the sensitivity of the statistical tests can be enhanced by:

(1) The use of non-parametric techniques on the gain score data wherein there is a strong likelihood of the lack of interval scaling and wherein extreme deviant performance has an inordinate influence upon the error term; and

(2) The use of a within S error term is a double classification, "mixed" ANOVA.

Both of these procedures were employed in further analysis of the results of certain of the pre-post criterion measures. Two criteria were employed in the selection of data to be subjected to additional analysis:

(1) Any case wherein there was a suggestion of superior performance upon the part of the control group;

(2) Those cases wherein there was an empirical and/or rational basis to anticipate effects related to the experimental conditions. The conduct of these additional analyses is justified in light of the exploratory nature of this study. However, significant findings which obtain in the analyses to be presented subsequently in this section must be viewed with circumspection. "Acceptance" of such results must be contingent upon the ability to replicate the findings in future experimental efforts.

Further Analysis of the Word Meaning Subtest of the Durrell-Sullivan Reading Capacity Test (DSRC-WM). This subtest was selected for further analysis because of the high mean gain on the part of the control group relative to the experimental groups (6.28 vs 1.86, 1.54, and 2.85 for the CC, MAV, TAV, and MAud groups respectively). The word meaning subtest is a test of listening vocabulary. As the experimental conditions involved intensive vocabulary training, the superior gain on the part of the control group warranted further study.

The data from the pre- and post-experimental administrations of the DSRC-WM were analyzed using a double-classification, repeated measures, unweighted means solution, ANOVA. The summary table is presented in Table 11.

The results of the analysis presented in Table 11 suggest that performance on the post-test was superior overall to that on the pre-test. The absence of experimental condition or interaction effects suggest there is no difference among the experimental groups and that the improvement obtains as a function of the general school program, maturation, practice and/or regression effects.

Non-parametric procedures were also employed to evaluate the significance of the DSRC-WM data. The Kruskal-Wallis one-way analysis of variance by ranks (Siegel, 1956) was employed on the gain score data as an analog of the F test. The trend at the 10 percent level was considered adequate to warrant probe tests for differences among means.

The Mann-Whitney U test (Siegel, 1956) was employed to test for significant differences among the four groups. The results suggested that the gain of the control group was superior to that of all the experimental groups. This difference obtained in spite of the fact that over 500 words were identified and defined during the course of the lessons presented to the experimental groups. However, there was no commonality among the recognition words employed in the lessons and the stimulus words employed on the DSRC-WM. Thus, it would seem, from the results of this analysis, that the classroom teachers had a greater tendency to teach words which were in common with the test than did the experimental teachers.

<u>Further Analysis of the Word Meaning Subtest of the Durrell-Sullivan Reading Achievement Test (DSRA-WM)</u>. This subtest was selected for

Table 11

Summary Table of ANOVA (Unweighted Means Solution)

on Pre- and Post-Data Collection for the

Four Groups on the DSRC-WM

Source	df	MS	F
Between Subjects			
A Exposure Conditions (CC, MAV, TAV, MAud)	3	202.3	2.0
Pooled Subjects within Groups	51	100.1	
Within Subjects			
B Test Administration (Pre-and Post-Test)	1	267.1	13.8***
AB	3	32.9	1.7
Pooled B x Subjects within Groups	51	19.4	

***F ratio significant at the 1% level

further analysis because of the high mean gain of the control group on the analogous test from the DSRC battery and the fact that a highly significant Fmax ratio (34.33) obtained which suggested that the use of the parameter-bound, simple randomized ANOVA was inappropriate. The Kruskal-Wallis one-way analysis of variance by ranks was employed on the gain score data from the DSRA-WM. The obtained value of H (uncorrected for ties) failed to reach significance at the 10 percent level; probe analyses for differences among individual groups were thus not appropriate.

Further Analysis of the Bond-Clymer-Hoyt Diagnostic Reading Test (BCH). Certain sections of the Bond-Clymer-Hoyt Diagnostic Reading Test were selected for further analysis. These were:

(1) Recognition Pattern (BCH-RP),
(2) Visual Analysis (BCH-VA),
(3) Phonetic Knowledge (BCH-PK).

(1) Further Analysis of the Recognition Pattern (BCH-RP) Gain Score Data. Dunnett's Test (Edwards, 1965) was employed to test for significant differences between each experimental group and the control group. The error term from the gain score, simple randomized ANOVA and the harmonic mean of the N for each group were employed to compute the standard error of the difference between two means. The results suggested no differences between the control group and any of the experimental groups. Non-parametric analyses confirmed the absence of differences among the four groups.

(2) Further Analysis of the Visual Analysis (BCH-VA) Gain Score Data. The Kruskal-Wallis one-way analysis of variance by ranks (Siegel, 1956) was employed on the gain score data from the Visual Analysis section of the BCH. It was con-

cluded that no significant differences exist among the groups as to gain on the BCH-VA.

(3) Further Analysis of the Phonetic Knowledge (BCH-PK) Data. Dunnett's Test was employed to test for significant differences between each experimental group and the control groups. The error term from the single randomized ANOVA of the gain score data and the harmonic mean of the N for each group were employed to compute the standard error of the difference between the means. The results of this analysis is presented in Table 12.

Table 12

Obtained T Values for Differences Between Control and Experimental Group Means on Gain on the Phonetic Knowledge Section of the Bond-Clymer-Hoyt Diagnostic Reading Scale

Group	$\bar{X}$ Exp.	$\bar{X}$ Control	$\bar{X}$ E - $\bar{X}$ C	T
MAV	12.45	1.73	10.72	2.34*
TAV	4.92	1.73	3.19	0.70
MAud	3.55	1.73	1.82	0.40

*T_{90}, df: 40 = 2.13

The use of the Dunnett Test reflects a difference in gain between the control and MAV group at the .10 two-tailed level. Non-parametric techniques were also employed to analyze the gain score data from the BCH-PK. A Mann-Whitney (Siegel, 1956) U test between the MAV and control groups yielded a U of 33.5 with a U of 37 or less needed for significance for a two-tailed test at the .05 level. Thus, these results suggest that the MAV group achieved a larger gain on the BCH-PK than did the control group.

Part II: Results of Daily Comprehension Tests Administered During the Instructional Phase of the Experimental Period.

There are three major sources of variance which might influence the error term used to evaluate the significance of comprehension as a function of exposure condition. These are variance associated with:

(1) The intrinsic difficulty of the passage content and the comprehension questions employed each day.

(2) The functional ability of the individual Ss to understand a passage of a given difficulty level.

(3) The interaction between passage (and question) difficulty and the comprehension ability of the children.

The above potential sources of variance were inherent in the experimental procedures employed in this study and, if extant, would influence any dependent criterion measure. In order to control for the influence of the above factors:

(1) Ss were selected so as to restrict the range of variance associated with listening comprehension.

(2) Passages were selected so as to be within the listening comprehension ability level of most, if not all, of the Ss.

(3) Estimates of readability were used to classify the difficulty level of the passages a priori.

A more serious concern was that of dealing with the rather extensive absenteeism on the part of the Ss. In order to deal with this problem, the basic criterion measure employed for inferential purposes was the percent of correct responses obtained by individual Ss over all days on which the S was present. This measure reflects the proportion of correct responses obtained by individual S in each group on the comprehension questions to which the S was exposed. In effect, this measure assigns his own mean performance score for any day the S was absent.

The percent of correct responses suffered from a number of limitations:

(1) There is a potential lack of sensitivity as a function of averaging performance over passages (and questions) of disparate difficulty and rate of presentation. One means of dealing with this is to classify the material according to difficulty and rate and to perform separate analyses. There should be a gain in sensitivity in such a procedure but the proportions involved would be based upon far fewer observations, implying a reduction in the reliability of the estimate.

(2) There is a potential source of bias in the percent of correct responses data. This bias is related to absenteeism:

(a) An interaction between S absence and the intrinsic difficulty of the passage and questions employed on the days the S was absent would tend to inflate or deflate the percent of correct responses.

(b) A difference in an S absence by passage difficulty interaction bias across the three experimental conditions would invalidate the percent correct response data as a dependent criterion index.

(3) Another limitation on the validity of the percent correct

response measure is the reliability of that index independent of potential bias associated with passage difficulty. Such reliability would vary as a function of the number of days S was present (the total number of observations which were employed in determination of the proportion.) The validity of estimates of the distribution of error variance associated with a lack of reliability would depend upon:

(a) The validity of assumptions as to the character of the sampling distribution of an inferential parameter.

(b) The adequacy of the obtained distribution of the criterion variable to serve as a basis for estimating parameter values.

(c) The position of a given point estimate as it influences the degree and direction of the distribution of error variance.

As one cannot assume homogeneity of variance in the case of proportional data, the arcsine transformation was recommended. However, this transformation required that proportions be based on the same N to assure homogeneity of variance. Since this was not the case, redundant analyses were performed upon both the non-transformed and arcsine transformed percent of correct response data in the case of analyses of central interest. However, these redundant analyses were consistent in outcome and, based on the sample variances, the assumptions of homogeneity of variance appeared tenable for the transformed data. Therefore, only the analyses on the transformed data are reported here.

An additional source of information concerning the influence of the experimental conditions upon comprehension is the mean of each experimental group on each daily test of comprehension. This measure is far more apt to be sensitive to the difficulty and rate of presentation of each passage than is the percent of correct responses for individual Ss taken over passages.

The daily group mean, however, also suffers from limitations:

(1) There is a lack of independence of each group's mean across passages and a lack of constancy of the group's daily constitution as a function of absenteeism. Thus, no valid statistical basis for inference can be established. The utility of the group mean data is limited to a descriptive function.

(2) The validity of the group mean data to serve a descriptive function depends upon freedom from the following sources of possible bias associated with absenteeism:

   (a) A difference in the relationship between Ss comprehension ability and the incidence of absences across the three experimental conditions.

   (b) A difference in an S-absence by passage difficulty interaction bias across the three experimental conditions.

Analysis of Data Based upon the Percent of Correct Responses Obtained by each Subject during the Daily Tests of Comprehension. These data are based upon the precent correct responses obtained by each S on tests of comprehension on that portion of the 30 of the 33 days which defined the instructional phase of the experimental period when the S was present. These 30 days are classified as follows:

(1) Those days upon which the groups were exposed to machine presentation of auditory materials under normal and compressed rates (6 days).

(2) Those days which involved expanded (very slow) presentation to the machine groups (24 days). These days are subdivided according to the difficulty of the passages employed:

   (a) Those days when easy material (2nd through 3rd grade level) was employed (12 days).

   (b) Those days when hard material (4th through 5th grade level) was employed (12 days).

Excluded from analysis ara data obtained on:

(1) Two days when field trips seriously influenced the constitution and size of the experimental groups.

(2) One day when a passage of high, disparate difficulty was employed under the expanded rate condition.

As the development of material proceeded concurrently with the experimentation the hard, easy, and compressed passages described above were presented in an interspersed rather than consecutive fashion. The actual distribution of these passages across time is given in Appendix C.

<u>Analysis of Arcsine Transformed Percent of Correct Responses by Individual Ss on Comprehension Measures Administered During the Instructional Phase of Experiment II.</u> Combined data covering all of the three divisions of the instruction phase (30 days) were analyzed using a simple randomized ANOVA to test the overall effects of the experimental conditions. Summary data and the results of analysis are presented in Table 13.

The results of the ANOVA summarized in Table 13 show an F ratio significant at the 10 percent level of significance. The Hartley Fmax ratio (1.65) was not significant (Fmax .99 w/k = 3; N = 30 = 3.3).

Duncan's New Multiple Range Test (Duncan MRT) (Edwards, 1965)

Table 13

Summary Data on the Arcsine Transformed Percent of Correct Responses on Comprehension Measures Administered During the Instructional Phase of Experiment II (N, each group = 15)

	Raw Data		
	MAV	TAV	MAud
$\bar{X}$	55.6	53.2	49.0
s	8.5	7.2	9.3

	ANOVA		
Source	df	MS	F
A: Exposure Condition	2	169.31	2.42*
Error	42	70.00	
Total	44		

*$F_{.90}$ (2, 42) = 2.42

for multiple comparisons was employed to test for the significance of differences in arcsine transformed percent of correct responses among the exposure conditions. The results of these probe analyses showed a significant difference (5% level) in favor of the mean MAV score as compared to the mean MAud score.

The data presented above suggest that the comprehension of the group exposed to bimodal stimulation with the machine was superior to that exposed to unimodal stimulation only, while the comprehension of the group exposed to bimodal stimulation with the teacher did not differ significantly from that exposed to unimodal stimulation only.

Analysis of Arcsine Transformed Percent of Correct Responses when Auditory Materials were Presented at Normal or Compressed Rates During the Instructional Phase of Experiment II. Data from the six days upon which auditory materials were presented at normal and compressed rates during the instructional phase were analyzed separately using a simple randomized ANOVA to test the effects of the experimental conditions. Summary data and the results of the analysis are presented in Table 14.

The results of the analysis presented in Table 14, although in the same direction as before, failed to reach significance. The Fmax ratio (1.55) was non-significant.

Analysis of Percent Correct Response Data Gathered under Condition of Expanded Presentation. Summary data for the arcsine transformed percent of correct responses for the expanded condition are reported in Table 15, separately for:

(1) All 24 days of expanded presentation upon which material ranging from 2nd to 5th grade level in difficulty were employed.

Table 14

Summary Data on the Arcsine Transformed Percent of Correct Responses on Comprehension Measures Administered when Auditory Materials were Presented at Normal or Compressed Rates

	Raw Data		
	MAV	TAV	MAud
$\bar{X}$	59.7	58.1	53.7
s	10.6	11.4	13.2
N	15	13*	15

	ANOVA		
Source	df	MS	F
A: Exposure Condition	2	143.58	1.025
Error	42	140.12	
Total	44	283.70	

*2 Ss were not present on any of the six days wherein normal and compressed rates were employed.

Table 15

Summary Raw Data on Arcsine Transformed Percent of Correct Responses Obtained under Expanded Presentation During the Instructional Phase of Experiment II

	MAV	TAV	MAud
24 Days of Expanded Presentation			
$\bar{X}$	55.0	52.1	48.0
s	8.8	7.4	8.4
12 Days of Easy Material Presented under Expansion			
$\bar{X}$	57.1	55.7	50.2
s	9.3	10.5	9.7
12 Days of Hard Material Presented under Expansion			
$\bar{X}$	51.3	48.4	44.9
s	8.9	6.5	8.4

(2) The 12 days upon which easy material (2nd to 3rd grade) were employed.

(3) The 12 days upon which hard material (4th to 5th grade) were employed.

The arcsine transformed percent of correct responses for the easy and hard material presented under expansion was analyzed using a double classification, repeated measure ANOVA. The independent variables were: (a) exposure condition (MAV, TAV, MAud); and (b) level of passage difficulty (easy, hard). The summary table for these analyses is given as Table 16.

The results of this analysis reflect a main effect of difficulty on comprehension of all groups, with the easy (2nd-3rd grade) material superior to that of the hard (4th-5th grade) content. Although the F-test did not quite reach the 10% level, the Duncan MRT was employed to probe for differences among the means of the exposure condition groups. The arcsine transformed percent of correct responses was tested with separate error terms for each difficulty level pooled across exposure conditions. The differences among ordered means showed that the MAV mean exceeded the MAud mean at the 10% level for easy material and at the 5% level for hard material.

Group Performance Means on the Daily Tests of Comprehension. The group daily mean data are organized in much the same way as the percent correct response data above. As no adequate model for statistical inference was available, graphic presentation of the data and qualitative commentary as to implications are employed. It should be remembered that these means are based on 10-item tests. Summary data of group daily mean performance are presented in Table 17.

Table 16

Summary of the Results of ANOVA of the Arcsine Transformed Percent of Correct Responses Obtained on Comprehension Tests Over Easy and Hard Material Presented Under Expansion

Source	df	MS	F
Between Ss			
A: Exposure Condition	2	318.99	2.26
Ss w/ Groups	42	141.13	
Within Ss			
B: Difficulty Level	1	796.65	39.49***
AB	2	33.45	1.65
B x Ss w/Groups	42	20.17	

$F_{.90}(2, 42) = 2.42$

***$F_{.99}(1, 42) = 7.28$

Table 17

Summary Data of Group Mean Performance on Daily Tests of Comprehension During the Instructional Phase of Experiment II

	Experimental Condition and Group		
	MAV	TAV	MAud
6 Days of Normal or Compressed Presentation			
$\bar{X}$	7.2	7.1	6.4
s	.67	.78	.69
24 Days of Expanded Presentation			
$\bar{X}$	6.6	6.0	5.4
s	1.3	1.5	1.2
12 Days of Easy Expanded Material			
$\bar{X}$	7.0	6.4	5.8
s	1.2	1.4	1.2
12 Days of Hard Expanded Material			
$\bar{X}$	6.1	5.6	4.9
s	1.2	1.4	.98

Figures 4, 5, and 6 show plots of the daily means for compressed/ normal, hard expanded, and easy expanded material, respectively. The plots are arranged in chronological order with the actual day of administration shown on the baseline. In general these figures show a consistent trend for the MAV group to surpass the MAud group and a tendency for the TAV group to lie between.

The composite trend of these graphs across time fails to show a clear-cut improvement, although the movement of the composite trend obviously exceeds the variation of the groups around the composite trend. The trend with time is probably affected by differential difficulty, motivation, and interest in the various passages. Overall, however, it appears that performance tended to improve with time during the first 18 to 20 days of the experimental period, falling off sharply for the next 6 to 8 days and recovering thereafter. The exact reason for such a decline is not entirely clear, but is of a magnitude that is hardly ascribable to chance factor.

One factor which was undoubtedly at least partially responsible for this dip was the fact that the Project Director was forced to increase the speed and difficulty of the presentations somewhat in order to avoid ceiling effects on the 10-item tests. These adjustments were made on a subjective basis.

One of the first questions that must be raised in evaluating the group daily mean data particularly with respect to interpreting the direction and degree of differences between exposure condition means on any given day, is the degree to which the relative difficulty of passages was constant for all groups. Another way of stating this question is to ask if there is an interaction between experimental condition and passage difficulty. Such an interaction could obtain as a function of a direct influence of the independent variables on an experimental group's ability

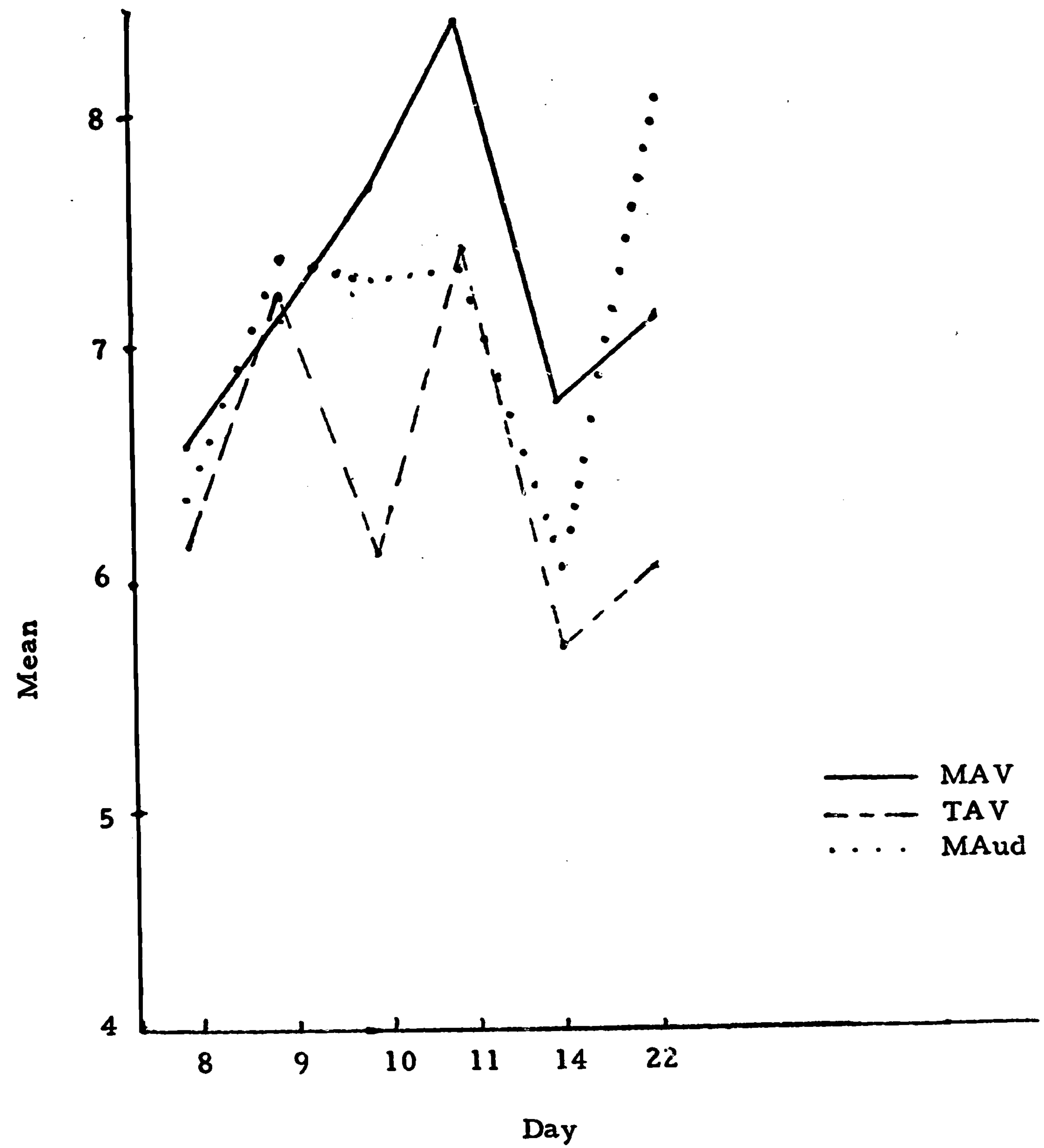


Fig. 4 Daily means for 6 days of normal/compressed presentation.

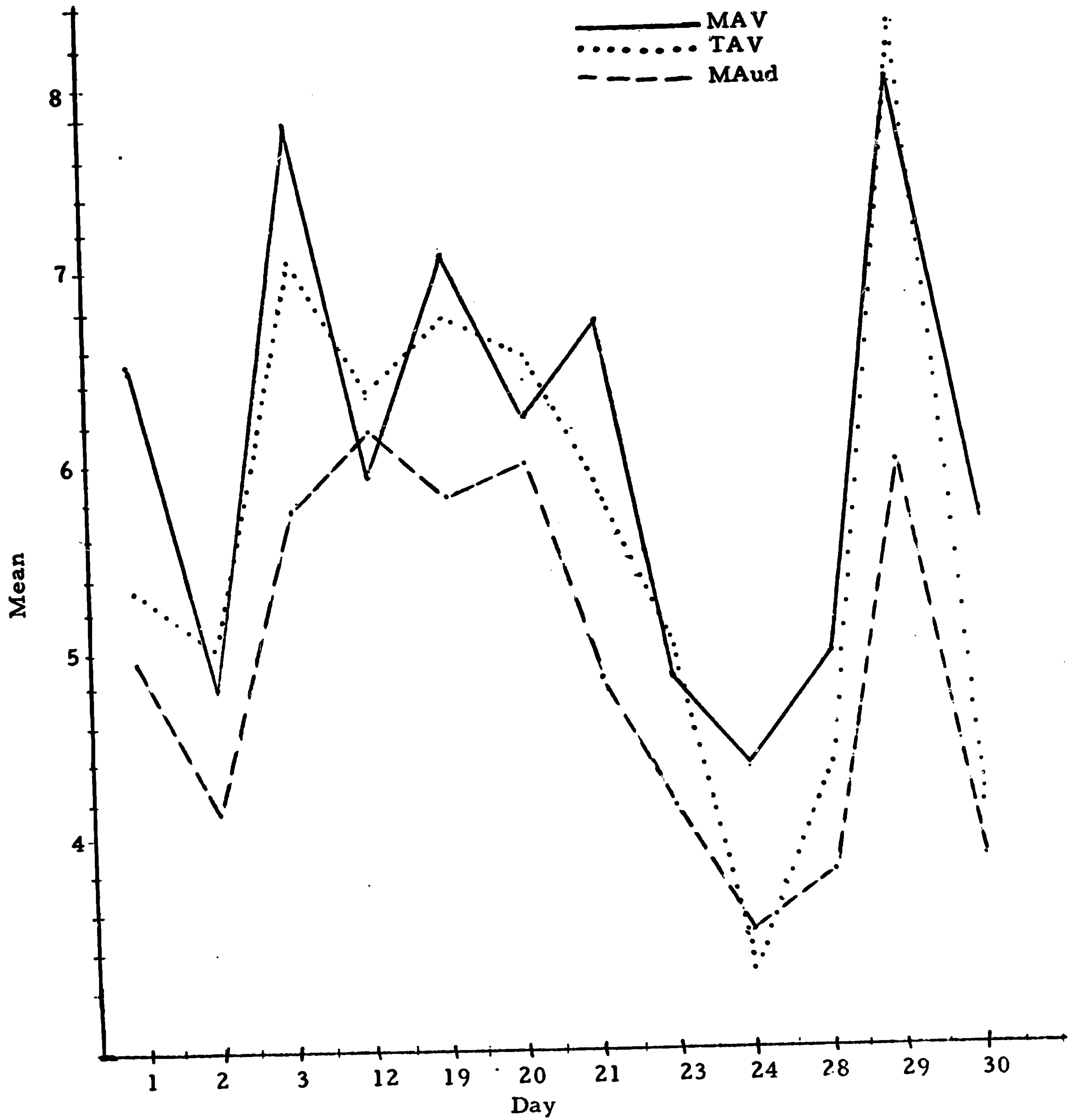


Fig. 5 Daily means for 12 days of hard, expanded presentation.

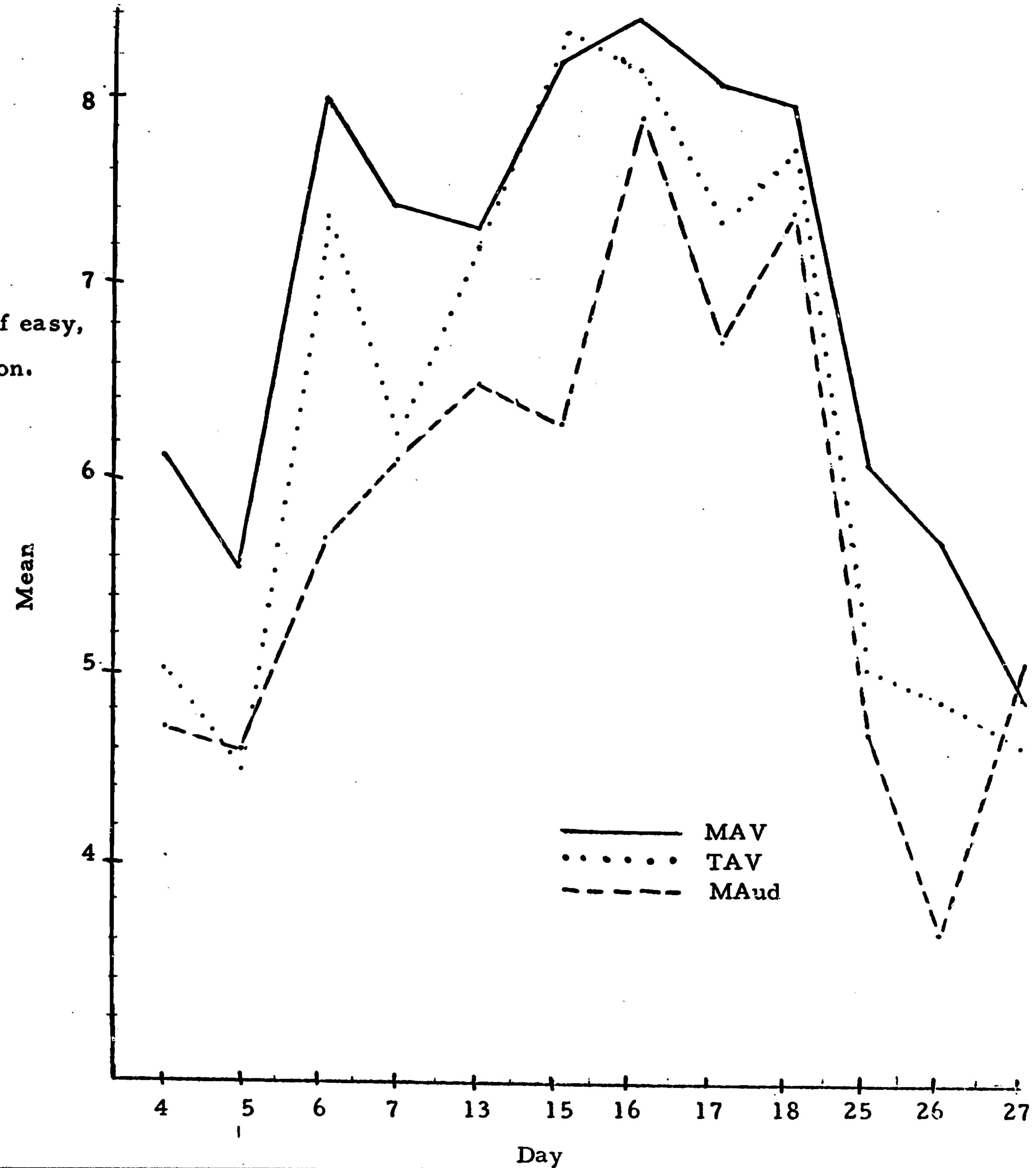


Fig. 6 Daily means for 12 days of easy, expanded presentation.

to understand a passage of a given difficulty level or, as mentioned previously, a difference in an S absence by passage difficulty interaction bias across the three experimental conditions.

The means of the three exposure conditions on the 12 days of expanded presentation of easy material, the 12 days of expanded presentation of easy material, the 12 days of expanded presentation of hard material, and all 24 days of expanded material (2nd through 5th grade difficulty) were analyzed using the Kendall Coefficient of Concordance (W). According to Siegel (1956), W varies as a linear function of the average Spearman rank order correlation (R_s) taken over all groups. In the present context, W represents the degree to which the relative difficulty of the passages was constant for the three exposure condition groups. The values of W and the average R_s associated with each set of data are as follows:

Source	W	Average R_s
Easy Expanded Material	.98	.97
Hard Expanded Material	.95	.93
All Expanded Material	.92	.88

The high degree of concordance among the ranked means of the exposure groups over all passages suggests little disparity as to the relative difficulty of each passage for each exposure condition.

This concordance is assumed to be primarily a function of the intrinsic difficulty of the passage (and questions) but would include practice effects as well. There seems to be little room for a source of bias related to an interaction between experimental condition and passage difficulty. Thus, whatever impact absenteeism might have had upon the group performance, it was not so extreme as to influence the performance of a given group on a given passage relative to other passages.

Siegel (1956) cites Kendall (1948) as suggesting that the order of the sum of ranks for each of N objects (in this case, passages) is the best estimate of the "true" ranking of the objects. According to this rationale, the sum of the ranks of means across all conditions for each passage is the best estimate of the relative difficulty of those passages (and questions). The Spearman rank order correlation between the sum of the ranks of means across experimental groups for each passage as an independent estimate of readability (the average Dale-Chall or Spache readability score on at least three samples) presented under expansion was +.05. It is clear that estimates of readability had little utility, in the present context, in predicting comprehension scores.

More prior research studies have yielded low but significant positive correlations between indices of readability and listening comprehension. In the present case, a primary determinant of the relative values of group means may have been the difficulty of the questions and not the passages. Furthermore, the reading measure may have been inappropriate for this population in this situation, particularly since the primary determinant of the readability estimate is vocabulary load. In the present study, the children were taught the words they were not apt to know.

Since the possible differential effect of treatment condition as a function of difficulty was of interest, and since the readability estimate bore little relationship to mean score (vide supra) the summed ranks of daily means was employed as an operational index of passage difficulty. Therefore, in the subsequent plots of the group daily means, the passages have been reorganized according to the summed ranks of the daily means (low to high ranks, difficult to easy passages), and the treatment means plotted. Figures 7, 8, and 9 present these means for the material ori- designated as easy expanded, hard expanded, and compressed/normal.

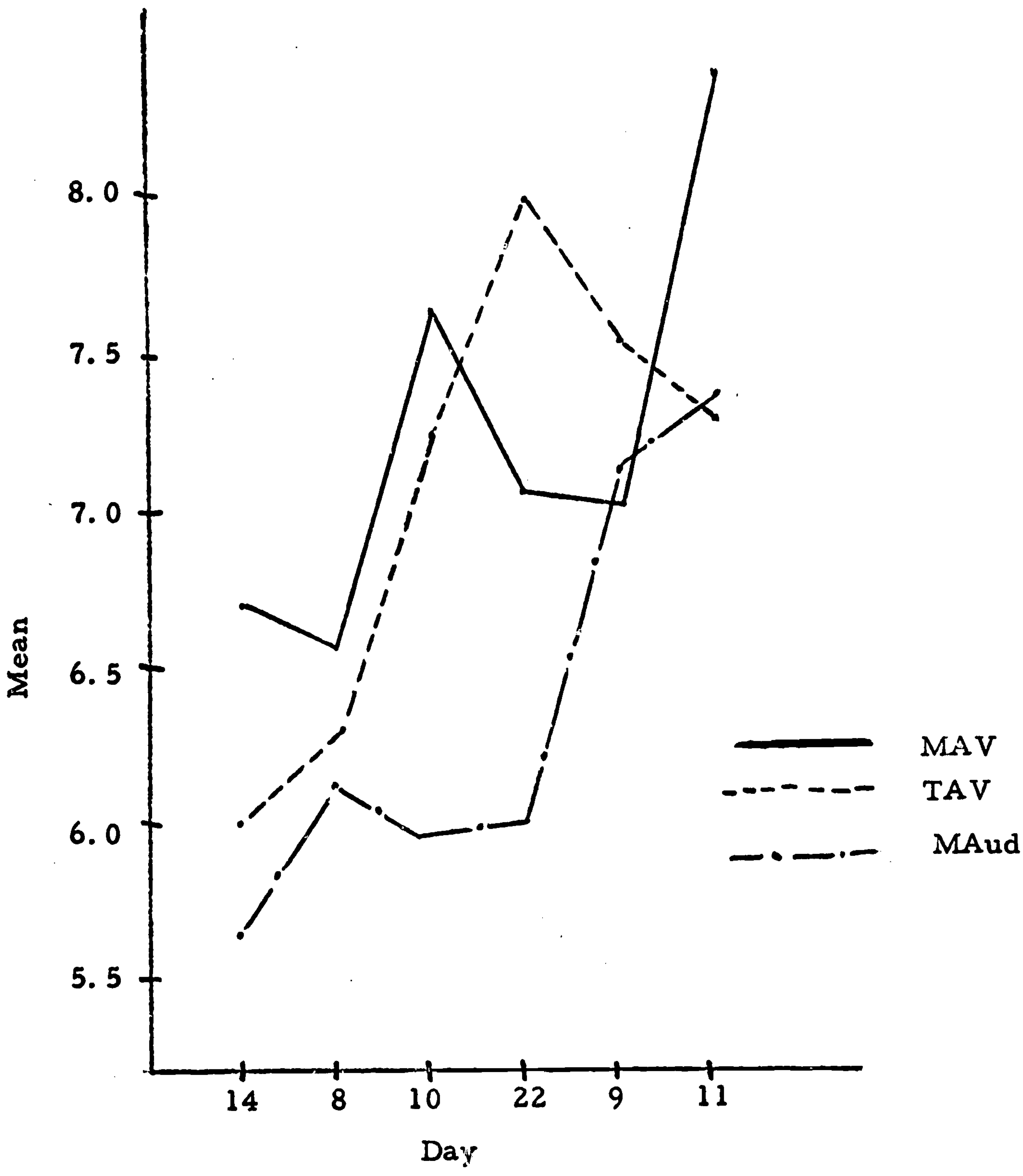


Fig. 7 Daily means for 6 days of normal/compressed presentation, in order of difficulty.

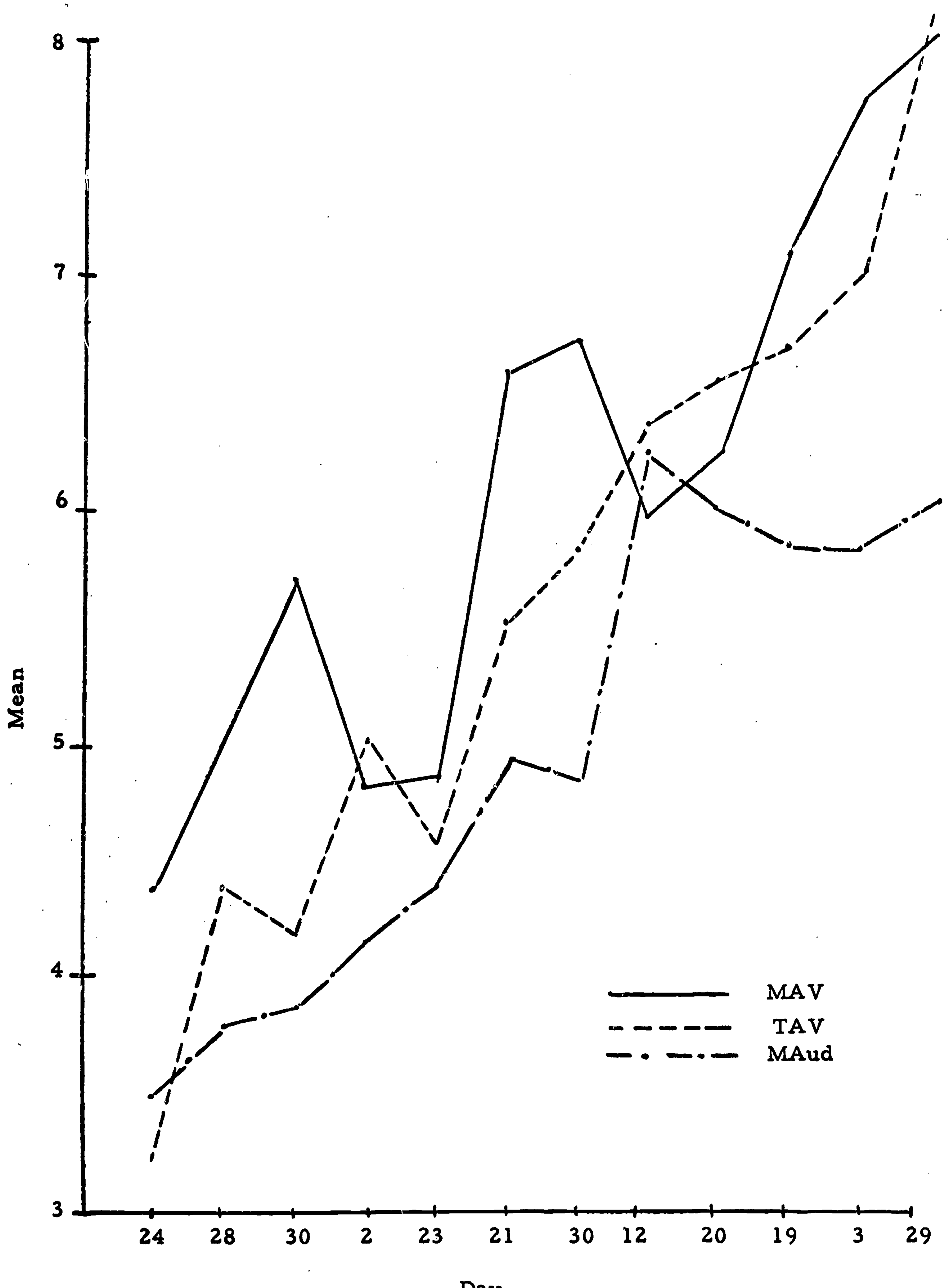


Fig. 8 Daily means for 12 days of hard, expanded presentation, in order of difficulty.

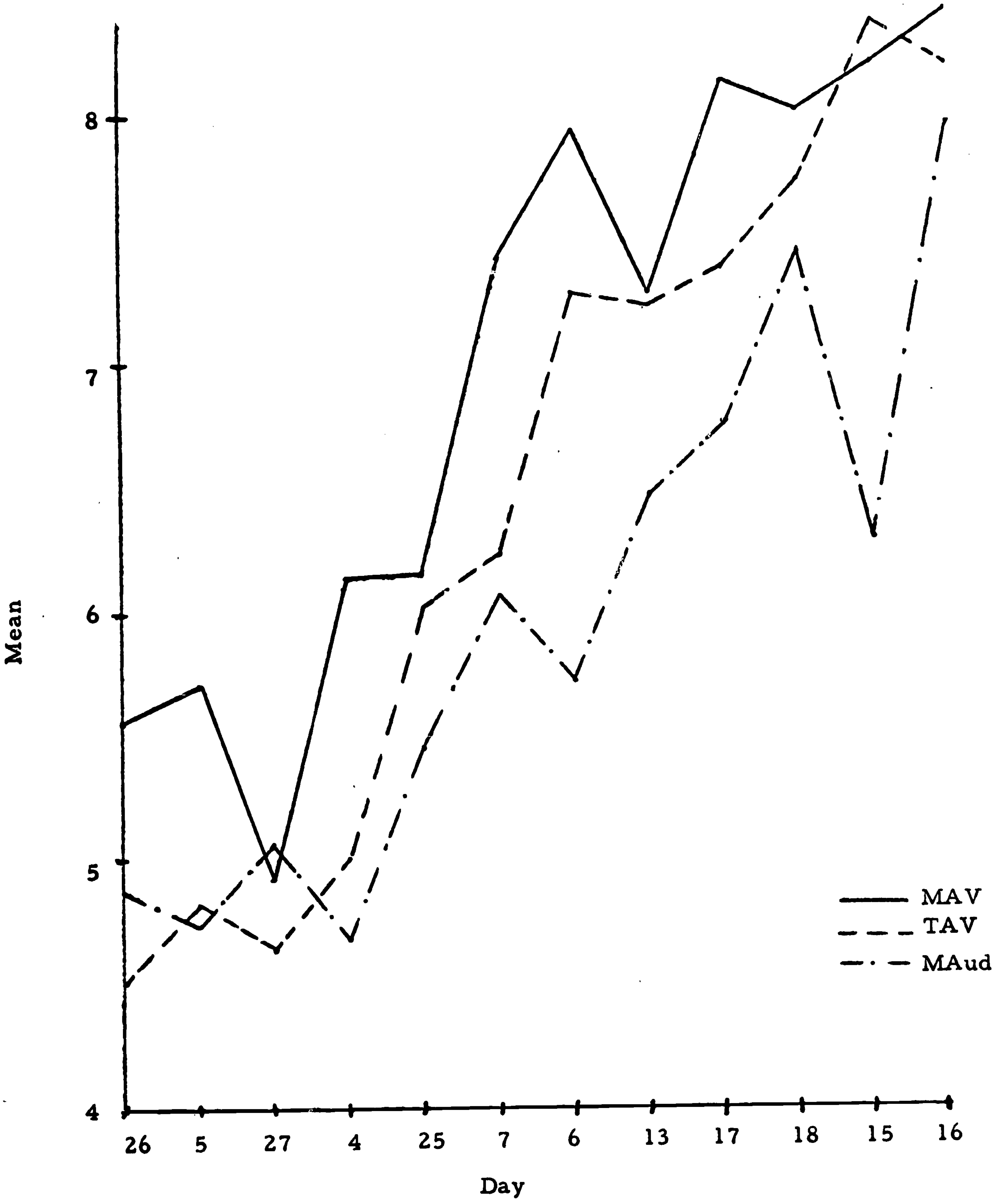


Fig. 9 Daily means for 12 days of easy, expanded presentation, in order of difficulty.

These figures again suggest rather consistently that the performance of the MAV group exceeded that of the TAV which exceeded that of the MAud group. For the expanded material (Figures 8 and 9) the TAV group appeared to be closer to the MAud group on hard material, but closer to the MAV group on easy material. No other relationship with empirical difficulty was apparent.

Interest Ratings. It will be recalled that the Ss were asked to rate the passages for interest on an ascending scale from 1 to 5. Conceivably some of the results obtained might have been attributable to this variable. Analyses of variance of the interest ratings data yielded a treatment group difference significant at the .01 level. Mean interest ratings were: MAud 3.33; MAV 3.99; TAV 4.30. A Duncan test indicated that the TAV and the MAV mean ratings were higher than MAud mean ratings at or beyond the 5% level of significance, while TAV and MAV group ratings did not differ. No differences in interest ratings by difficulty level were detected. These data fail to support the hypothesis that the trend for the MAV performance to be superior the TAV performance may be attributed to a higher level of interest in the material. Of course, the interest ratings were a product both of material and presentation mode. Therefore, it appeared that the bimodal presentation modes produced higher interest levels than did the audio only presentation mode.

Part III: Analysis of Data from the Retention Phase of Experiment II

It will be recalled that three special passages (of difficulty Grades 3, 5, and 9, respectively) were administered in that order to all treatment groups at the very beginning of the experiment. They were presented in accordance with the type of treatment prescribed for the groups (MAV, TAV, or MAud). Subsequently the questions only were given to a control group. As before, all tests were given using both the reading and listening modes. Group data collection efforts were conducted on nine days. Make-up sessions for absent Ss were limited to those possible to conduct within one calendar week of the group data collection effort as the influence of long-term temporal intervals between the initial, retention, and relearning stages was a variable of interest. As a consequence, there was a differential N in each group, and thus, the potential influence of a "self-selection" bias associated with Ss dropped from this phase may limit the interpretation of these data.

Tests showed in general that the treatment group means significantly exceeded the control group mean, thus indicating, as expected, that the scores were a function of knowledge gained from the passages. (It will be remembered that the control group did not receive the passages.)

As measures of retention, the "odd" items of the tests were re-administered one month after the initial data collection, and the "even" items were re-administered two months after the initial data collection. Of course, it is desirable that the "odds" subtests not differ from the "evens" subtests in making these comparisons. Therefore, this question was tested by analyzing the mean "odds" versus the mean "evens" scores by treatment group and passage, for the initial administration data. The appropriate means are shown in Table 18. In order to simplify the analysis of variance, the Ns were equalized in each cell to an N=8 by randomly casting out the excess cases.

The ANOVA of these data is summarized in Table 19. This analysis clearly shows that the two subtests (odds and evens) were not different, at least with respect to initial score data. It may be safely presumed that they were not different for other administrations, and the absence of interaction between split and treatment suggest that the two subtests were appropriate for the comparison of treatments.

Table 18

Means and Standard Deviations by Odd and Even Items, Treatment Groups, and Passages for Initial Comprehension Data, Retention Phase

Difficulty (Passage)		Control		MAV		TAV		MAud	
		Even	Odd	Even	Odd	Even	Odd	Even	Odd
3	($\bar{X}$)	3.8	3.6	7.5	7.6	5.2	5.8	5.8	6.0
	(s)	.88	.92	2.07	2.07	1.28	1.39	1.91	2.27
5	($\bar{X}$)	3.3	3.3	6.6	5.7	6.0	5.1	4.0	4.2
	(s)	1.17	.88	1.60	1.98	1.60	2.30	2.07	2.19
9	($\bar{X}$)	2.0	3.0	5.1	5.1	4.6	5.2	3.0	3.1
	(s)	.93	.75	2.10	1.46	2.26	2.66	2.45	2.03

Note.--Each entry is based on N=8; data is based on 10-item tests. The Control group did not receive the passages and thus represents a baseline estimate.

Table 19

Summary of ANOVA for Initial Data Retention Phase, Experiment II, by Odd-Even Split, Difficulty and Treatment Group

Source	df	Mean Square	F
Between Ss			
A Treatments	2	45.05	2.97*
Ss Within Groups	21	15.14	
Within Ss			
B Odd vs. Even Split	1	.01	-
A x B	2	.067	-
B x Ss Within Groups	21	2.26	
C Difficulty	2	45.09	19.0***
A x C	4	6.33	2.67**
C x Ss Within Groups	42	2.37	
B x C	2	2.43	2.05
A x B x C		.80	-
BC x Ss Within Groups	42	1.21	

*p < .10

**p < .05

***p < .01

As expected, the Grade 3, 5, and 9 passages proved highly detectably different in difficulty. However, a treatment effect was also detected at the 10% level with MAV> TAV >MAud being the order of the means. Further analyses of the treatment effect were suggested by the treatment by difficulty interaction. Analyses of the simple main effects confirmed the treatment difference at the 10% level or better for each of the three difficulty levels. The means were in the order given above for the Grade 5 and Grade 9 passages and MAV> MAud> TAV for the Grade 3. Not all mean differences reached significance however, the significant ones were as follows: Grade 3, MAV> TAV, MAud; Grade 5, MAV> MAud; Grade 9, MAV, TAV> MAud.

Having demonstrated that the odd-even subtests were equivalent, the retention data themselves could then be analyzed. The average of the odd-even scores was taken at the initial stage and compared to the one month and two month scores by initial treatment group, by difficulty (passage). Thus, these analyses were based on 10-item tests or equivalent. The means are shown in Table 20. The purpose of this comparison was to determine whether or not there were significant differences in score after the two intervals of time, and whether or not difficulty and initial treatment differences persisted.

The ANOVA is summarized in Table 21. It will be noted that differences in initial treatment did not hold up over the retention intervals, although the overall differences in the means remained in the same direction. Mean differences by difficulty level were again highly significant. Mean differences at the retention intervals were also significant and in the anticipated direction: initial > one month> 2 months. These analyses suggested that the favorable effects of bimodal presentation were decreased over the retention intervals, but that the bimodal procedure did not adversely affect retention. In other words, the results may be interpreted as a case of the regression of the high groups toward the common mean, across time.

Table 20

Means and Standard Deviations by Retention Interval, Difficulty and Treatment Groups (Cell Ns=8)

Difficulty (Passage)		MAV R_0	MAV R_1	MAV R_2	TAV R_0	TAV R_1	TAV R_2	MAud R_0	MAud R_1	MAud R_2
3	($\bar{X}$)	7.6	6.5	6.4	5.5	6.0	4.8	5.9	5.8	4.0
	(s)	1.76	1.14	2.26	1.20	1.85	1.28	1.90	1.28	1.85
5	($\bar{X}$)	6.2	5.2	5.0	5.6	5.9	4.9	4.1	4.8	4.6
	(s)	1.42	2.19	2.20	1.68	2.95	1.73	1.86	1.28	2.82
9	($\bar{X}$)	5.1	4.4	4.1	4.9	4.0	3.8	3.1	3.6	3.2
	(s)	1.73	1.51	1.36	2.38	2.67	1.49	2.13	1.41	1.58

Note.--Data is based on 10-item tests, where R_0 is the average of initial odd and initial even scores.

Table 21

Summary of ANOVA by Retention Interval,

Difficulty, and Treatment Group

Source	df	Mean Square	F
Between Ss			
A Treatment	2	29.14	1.90
Ss Within Groups	21	15.34	
Within Ss			
B Retention Interval	2	12.42	4.33**
A x B	4	2.70	-
B x Ss Within Groups	42	2.87	
C Difficulty	2	58.49	18.6***
A x C	4	3.14	1.24
C x Ss Within Groups	42	2.53	
B x C	4	1.48	1.01
A x B x C	8	1.69	1.15
BC x Ss Within Groups	84	1.47	

**p < .05

***p < .01

A further measure of retention was obtained by re-administering the passages to the three treatment groups, using the original presentation mode, two months after initial exposure. The data are summarized in Table 22. The ANOVA (Table 23) showed significantly higher mean scores for the relearning condition than for the initial data. The ANOVA again revealed no differences by treatment group and highly significant differences by difficulty level in the expected direction. However, the nearly significant treatment by test interaction suggested that the initial vs. relearning increase be examined by treatment group. In an analysis of simple main effects this difference was found to be significant for TAV and MAud, but not for MAV. A similar analysis showed that while the means remained in the order MAV > TAV > MAud, the significant difference obtained on the initial test did not hold up on the relearning test.

Thus it would appear that the second administration of the passages produced scores at a significantly higher level than the original administration in spite of a two-month interval between administrations. It should be noted, however, that the two-month interval contained 30 instructional periods in which the experimental treatments were used with other materials. The demonstrated increments could thus be attributed to relearning or treatment effects, or probably, both. In any case, these findings suggest that EMRs can profit from learning experiences. The failure of the MAV increase to reach a significant level was again probably due to a regression effect of the highest score level toward the common mean.

A final question which may be investigated with these data is whether or not the scores obtained after an interval differed significantly from a chance level. Relevant information can be had by comparing the scores for the treatment groups after one month to those obtained by the control group (which did not receive the passages at all)

Table 22

Means and Standard Deviations for Relearning

by Difficulty and Treatment Group

Difficulty (Passage)		MAV		TAV		MAud	
		RL_0	RL	RL_0	RL	RL_0	RL
3	($\bar{X}$)	15.1	15.4	11.0	14.4	11.8	13.1
	(s)	3.52	4.00	2.39	2.50	3.81	4.09
5	($\bar{X}$)	12.4	13.2	11.1	13.1	8.2	11.9
	(s)	2.97	5.09	3.36	3.56	3.73	5.17
9	($\bar{X}$)	10.2	12.1	9.9	11.6	6.1	9.4
	(s)	3.45	4.82	4.76	4.53	4.26	4.41

Note.--Cell Ns=8; data based on 20-item tests.

Table 23

Summary of ANOVA for Initial and Relearning Scores

by Difficulty and Treatment Group

Source	df	Mean Square	F
Between Ss			
A Treatments	2	109.2	1.79
Ss Within Groups	21	60.9	
Within Ss			
B Relearning	1	150.06	34.9***
A x B	2	10.19	2.37
B x Ss Within Groups	21		
C Difficulty	2	152.30	18.22***
A x C	4	9.19	1.10
C x Ss Within Groups	42	8.36	
B x C	2	13.15	1.90
A x B x C	4	5.12	-
BC x Ss Within Groups	42	6.91	

***$p < .01$

for the same (odd) items. The relevant means are contained in Tables 18 and 20.

The ANOVA is shown in Table 24. A highly significant difference in the group means was detected. Examination of the means themselves showed that this difference arose because all treatment group means significantly exceed the control group mean, thus suggesting strongly that some of the material was retained by the treatment groups over the interval of one month. Again a significant difference in difficulty was noted.

Although statistical interdependencies precluded a test at the two-month interval, it seems safe to say that EMRs demonstrated a significant capacity for the retention of intructional material over a considerable time span. Means at the two-month interval were in the same direction. This finding may be one of the more important of the study as it indicates the value of continued effort to improve instructional techniques for the EMR.

To summarize the retention phase of the study, significant retention was obtained after a one-month interval (and probably after a two-month interval) as compared to a control baseline that did not receive the passages. Significant losses with time were also detected, as expected, by comparison with initial scores. The initial differences in the treatment group means in the direction MAV > TAV> MAud did not hold up across the retention interval. A significant improvement over initial score was found on a relearning exercise carried out after a two-month interval, particularly for the TAV and MAud groups. Difficulty levels were significant in all analyses in the expected direction. It was concluded that EMRs are capable of retaining and of significantly profiting by relearning for the instructional materials used; and that the initial treatment group differences tended to regress toward a common mean over time.

Table 24

Summary ANOVA of Treatment Means and Control Means at the One Month Retention Interval by Difficulty Level

Source	df	Mean Square	F
Between Ss			
A Groups	3	29.21	6.56***
Ss Within Groups	28	4.45	
Within Ss			
B Difficulty	2	32.85	14.0***
A x B	6	.62	-
B x Ss Within Groups	56	2.34	

***$p < .01$

Discussion and Conclusions

The results of Experiment I have already been discussed in conjunction with their presentation.

The findings of Experiment II are discussed below. In general, the 10 percent level of significance was employed in the analyses since the very small groups and very short tests employed suggested that Beta error might be more prevalent than Alpha error.

Pre-Post Test Data

It was hoped that the experimental procedures would result in a measurable increase in the reading skills of EMRs. In general, this hypothesis was not supported. In considering this finding, it should be noted that the total amount of exposure of the Ss to the experimental procedures was only about 30 hours, of which only about half was spent in instruction (the remainder being testing and administrative). It may very well be that significant changes (measurable with instruments not fully appropriate) cannot be expected in so short a time.

One difference which was detected was with respect to improved phonetic knowledge for the MAV group. This is a very reasonable finding, considering the bimodal presentation at a rate slow enough to enable the EMR to match sight and sound.

Daily Comprehension Tests

These data were analyzed both in terms of the percentage correct questions overall daily tests, and in terms of plot of the daily means for the three treatment groups.

Percentage Correct

The overall ANOVA for the 30 days of the Instructional phase indicated treatment differences at the 10 percent level. While it was only possible to statistically detect that the MAV mean exceeded the MAud mean, using probe analyses, the means were in the order MAV > TAV > MAud -- an order which occurred repeatedly in various analyses and across various levels of material difficulty throughout Experiment II. When the data were subdivided into the three categories of normal/compressed rate; expanded rate, easy material; and expanded rate, difficult material; it was no longer possible to detect treatment differences, although the F for the expanded material approached the 10 percent level and the means consistently maintained the direction MAV > TAV > MAud. The authors are inclined to feel that treatment differences may have existed, but that the tests lacked the power to detect them statistically.

Daily Means

The plots of the daily means tended to support the above conclusion, in that rather consistent trends were observed.

Retention Data

The retention data also tended to support the general conclusion that the treatment groups differed, again in the same direction. This important support came from entirely different data, and after no prior exposure to the experimental treatment (see the analysis of initial retention data). This finding, in conjunction with the observation that group differences did not appear to be correlated with time, suggests that any treatment effects to be expected may occur almost immediately, and that greater exposure than employed here may be required to produce a practice effect.

Further, the retention data indicated strongly that the EMR has the capacity to profit from the kinds of instruction studied here.

Discussion of the Findings

Several sources of variance could have been involved in the relative trends (of degree and direction) among the comprehension data for the three exposure condition groups. Consistent trends could obtain as a function of:

(1) Differences as a function of the mode of presentation (treatment). This, of course, is the dimension of interest. The problem is that other sources could also generate such differences as might obtain. No definitive basis for rejecting the alternatives exists, but a rational argument can be made that the treatment modes were the major source of the differences (see below);

(2) There might be differences which inhere in the Ss themselves. This, of course, is a serious possibility associated with any between Ss comparison. If such a source of bias existed, however, it was unique to the daily comprehension measure as no differences among the groups obtained upon the pre-experimental administration of the extensive battery of standardized tests;

(3) The group means obtained on a given day would vary, to a certain extent, as a function of the comprehension skills of those Ss who were absent on a given day. The influence of such a bias would vary as a function of the difficulty of the passage and questions employed on that day. The impact of such bias might influence the relative position of each group on a given day or the position of a given group's daily mean relative to performance by that group on other days. However, there was a high degree of concordance, as obtained by the Kendall Coefficient of Concordance (W), among the means as to the relative difficulty of each passage. This suggested that it is valid to interpret the relative position of each group mean on a day-by-day basis;

(4) Another source of potential bias would be a difference among the three groups in the degree to which Ss of disparate ability displayed a differential incidence of absenteeism. One source of information as to this possibility was the correlation between ability and attendance of the Ss in each group. The Spearman rank correlations between attendance and the

DSRC total score and the PMA total raw score suggested that no such bias played a significant role in the determination of the daily means of the three groups.

In line with the above argument, the comprehension data can be interpreted under the assumption that the differences among the means were a function of the exposure condition rather than bias associated with S ability and/or absenteeism. The plots of the group daily means supported the conclusions based upon the percent correct response data. These plots also suggested that the comprehension of the MAV group was superior to that of the TAV group. In those 10 days whereupon the TAV group was superior to the MAV group, it was generally possible to offer, post hoc, a basis for why the machine group was inferior.

The general superiority of the MAV group was fully not anticipated and no adequate means of developing a basis for statistical inference was incorporated into the experimental procedures as this would have entailed the administration of many makeup sessions for absentees. It might have been assumed, a priori, that if there was any consistent trend in the comprehension daily mean data that it should favor the TAV group. The bases for such an assumption are as follows:

(a) The children had had years of experience with auditory materials presented at normal speaking rates and no prior exposure to the auditory presentation of connected discourse at extremely slow rates of speed;

(b) There was some distortion of vocal quality associated with expansion which could not be eliminated. It was assumed that the effect of such distortion should have negative effects:

(c) The test questions were presented by live speakers at "normal" rates. Thus, the commonality between the stimulus complex associated with passage presentation and test conditions was optional for the TAV group.

The finding that the MAV group appeared to be superior in comprehension to the TAV group in spite of having the deck stacked against them made the comprehension data the major locus of experimental interest rather than a control for possible loss to be considered in evaluating the use of the machine to increase specific word attack skills and reading vocabulary. This did not seem to be associated with differences in the Ss. Teacher-pupil rapport seemed to be, on a subjective basis, superior in the TAV group (S interest ratings tended to confirm this conclusion). The variance associated with the effect of machine versus live speaker presentation on comprehension was comparatively small relative to the effect of passage and question difficulty. On the other hand, anything that has the potential of increasing the comprehension by EMRs of instructional material is worthy of attention and further research.

No valid probability statement can be offered in support of the existence of a superiority in comprehension of the MAV group over the TAV group. However, based upon the consistency of this trend and the recognized limitation of small group experimentation, it can be assumed that a replication experiment would reflect the existence of such superiority. The following rationale is offered:

Given an assumed superiority on the part of the MAV (under expansion) over the TAV group in comprehension, what would account for the difference? These children had had years of exposure to verbal stimulation at "normal" speaking rates. Why should comprehension be better at rates substantially slower than that to which the Ss are habituated?

The most obvious alternative is that the enhanced performance was a function of novelty. However, anecdotal evidence supplied by teachers of the EMR militates against such an interpretation. The performance of EMRs is seldom enhanced, and is often disrupted by the introduction of novel stimuli. The interest rating data on the TAV bimodal presentations did not display a significant difference. The absolute values of interest ratings of the TAV group were higher than those of the MAV group. None of these observations precludes the existence of a novelty effect but it does seem unlikely that such was the case. A novelty effect interpretation does not seem to be of heuristic value in guiding future experimentation. It does, however, entail caution in suggesting alternative explanations of the data and must be considered in planning procedures for future experiments.

Since the performance of the MAV group seemed to be enhanced under conditions of expansion, the machine itself and/or the slow rate could be the basis for the superior performance. The machine did not enhance the performance of the MAud group relative to the TAV group but the appropriate control (teacher presentation of audio material only) was not available. It does seem clear, however, that the availability of the visual material was a necessary condition for enhanced comprehension associated with machine presentation of the audio stimuli.

Harris (1961, p. 510) suggests that "to some extent, rate of reading is related to rate of thinking. It does no good to try to read (or listen?) faster than one can assimilate ideas." A parsimonious, if rather simple-minded, interpretation of the data would be to assume that the "thought rate" of EMRs is slower than that of normals and that the slowing up of the auditory rate yields a better "match" with that rate. The comparatively low performance of the MAud group argues against such an interpretation, but does not preclude it, as a teacher-audio group might have done even less well. It would seem that the availability of the visual material is a necessary condition which must obtain before the EMR can "profit" from a slower rate of presentation.

Thus, it would appear that the results of the experimentation offer tentative support to the original hypothesis that simultaneous bimodal presentation of the same material would produce a facilitation effect as compared to unimodal presentation. While the role of the rate factor was not exactly clear, it is also probable that some slowing of rate below normal is desirable.

Additional experimental work would be highly desirable to clarify the findings with respect to rate; to examine further the interactions between rate and difficulty obtained in Experiment I; to reconfirm and further elaborate the conditions under which bimodality facilitates comprehension; and, using specially developed measures, to re-examine any possible effects of the experimental procedures on the improvement of reading skills in the EMR.

It is felt that the overall results of the exploratory work presented here have been promising. The paucity of effective instructional techniques for the EMR suggests that follow-up work along these lines would meet an important need. Furthermore, such additional experimentation would greatly extend the contribution of the present study toward the basic understanding of facilitative effects in human information processing and comprehension.

Appendix A*

Today's lesson is about:

JET AIRPLANES

We will learn:

What makes jet airplanes go.

What makes propeller planes go.

Here are some words we must know when we see them:

roar	understand	thickest
propeller	force	thrust
knife	splashes	kickback
engine	skyrocket	kerosene
oxygen	forward	balloon

Here are some words we must understand:

Oxygen - Oxygen is a gas that has no color or smell. It is found in the air we breathe.

Examples: Animals cannot live without oxygen.

Fires will not burn without oxygen.

Kerosene - Kerosene is a thin oil. It can be used as a fuel.

* Estimated readability level is Grade 5

JET AIRPLANES

You may have watched an airplane roar down the runway for a takeoff. It goes faster and faster. Then it goes up into the air. It goes higher and higher.

What makes the plane go up?

Why does it stay up in the air? It stays up because the air holds it up.

To understand how air works, you have to know what it is like. Sometimes air seems like nothing at all. This is because we can't see it. But it is both heavy and strong.

How much does air weigh? The air inside an empty suitcase weighs about an ounce. That is what a silver dollar weighs.

Air is all around us. It fills the space around the earth that does not have something else in it. The earth has a thick blanket of air all around it. The thickest part of this air blanket goes up about 20 miles into the sky.

But the air gets much thinner away from the earth. There is less air at the top of a mountain than at the bottom. Twenty miles up there is hardly any air.

You have a blanket of air above you. Then why don't you feel it pressing on you? The reason is that air presses in all directions.

How is an airplane lifted from the ground? The air must be made to hold up the wings. The plane will stay up if the air below the wings pushes up harder than the air above pushes down. But air usually pushes just as hard both up and down. How can we give it more push upward?

Air that moves acts differently from still air. When air moves fast it can't press down very hard. The air above the wing must move faster than the air under the wing. That way the air will have more push upward. The propeller pushes air both over and under the wings of the airplane. But the top of the wing is rounded. The air going over the top of the wing has to go farther than the air that goes under it. So the air must go faster to reach the back edge of the wing at the same time as the air which goes under it. The top air goes faster and farther. It cannot press down so much. It has less push than the slower air under the wing. So the air under the wing can give the wing a lift. An airplane also gets a lift from its speed. The faster the plane goes the greater will be the lift.

Propellers make the airplane go. As it turns, the propeller makes a force called thrust.

Just how does a propeller work? It has thin edges like a knife or an ax. A knife cuts by opening a way into something. Then the knife pushes the pieces apart. A propeller works like a knife. It keeps cutting into the air and pushing it back.

When the air is pushed back, it gives a little kick forward. This is the force called thrust. This is what makes the airplane go forward.

You can see how this works if you drop a stone into a pond. The stone goes into the water and pushes it down. But the water kicks back. It splashes up.

The kickback of the air against the propeller gives the propeller a push forward. This makes the plane go forward.

Most small airplanes have only one propeller. Many large planes have four propellers.

Planes that have no propellers are called jet planes. They get the thrust that pushes the plane forward from a jet engine.

A jet plane goes much faster than an airplane with a propeller. Some Air Force jets fly as fast as 2200 miles an hour. Some big jet airliners go about six hundred miles an hour.

A jet engine works like a Fourth of July skyrocket. The great big rockets we are shooting into outer space work like skyrockets, too. The plane is pushed forward by the hot gases that shoot out from the back.

The gases are made by burning kerosene or some other fuel. Nothing can burn without oxygen. The jet engine gets its oxygen from the air. Air comes into the jet engine from an opening in front.

How can just hot gases make a plane go forward? You can see how if you blow up a toy balloon. Hold it and let the air out. It blows out. You can feel the jet of air rushing out. Blow up the balloon again. Let go of it.

It will push away from you It will fall to the ground when it is empty.

The jet of air pushes out at the back. The balloon is pushed forward. The hot gases in a jet plane are much stronger than the air you put in a balloon. The jet of hot gases rushes back from the engine. The plane is kicked forward.

Jet planes were first made for use in war. Now they are used to carry anyone who wants to go in them. They go very fast.

With the jet plane getting pushed so fast, the wings get enough lift from the air to keep it up. It does not need a propeller. Jet planes can go much higher and faster than planes with propellers. Because of their great speed, the thin air high above the earth can give them enough lift to keep them up.

Watch a jet plane flying You can see long lines of gases coming out the back of the plane. These are what give the plane thrust. They push the plane through the air When we can see the gases they are no longer pushing the plane

TEST FOR: JET AIRPLANES

1. What holds an airplane up?

 ____________ A. Gravity

 ____________ B. The air

 ____________ C. The gas inside it

 ____________ D. The fast start it got when it took off

2. Air is thickest

 ____________ A. near the earth

 ____________ B. where the weather is hot

 ____________ C. in the mountains

 ____________ D. twenty miles up

3. Air usually presses

 ____________ A. just down

 ____________ B. just up

 ____________ C. only in the direction the wind is blowing

 ____________ D. in all directions

4. For an airplane to stay up, the air has to move fastest

______ A. in front of its wings

______ B. behind its wings

______ C. above its wings

______ D. below its wings

5. When air moves very fast it

______ A. moves upward

______ B. moves downward

______ C. pushes upward

______ D. doesn't push down as hard

6. Thrust is a

______ A. force

______ B. part of a plane

______ C. gas

______ D. kind of knife

7. How many propellers does a jet plane have?

____________A. One

____________B. Two

____________C. Four

____________D. None

8. Jet engines use kerosene and

____________A. hydrogen

____________B. helium

____________C. oxygen

____________D. nitrogen

9. One advantage of jet planes over planes with propellers is that jet planes can fly

____________A. lower

____________B. in all kinds of weather

____________C. faster

____________D. slower

10. What pushes a jet plane along?

____________A. Air pushing on the propeller

____________B. Gas pushing on the propeller

____________C. Air pushing upward on the bottom of its wings

____________D. Gases pushed from its back

APPENDIX B

Summary of Data Analysis
on Pre- and Post-Test Criterion Measures

Source		Pre-Test	Post-Test	Gain
Durrell-Sullivan Reading Achievement: Total Score (Grade Level)				
CC	$\bar{X}$	3.83	4.16	0.34
n=13	s	1.05	1.09	0.47
MAV	$\bar{X}$	4.07	4.34	0.25
n=15	s	1.02	1.28	0.84
TAV	$\bar{X}$	4.12	4.40	0.25
n=13	s	1.37	1.21	0.67
MAud	$\bar{X}$	4.05	4.52	0.48
n=13	s	0.85	0.91	0.46
Fmax		2.15	2.0	3.32
F		0.19	0.22	0.40
Durrell-Sullivan Reading Achievement: Total Score (Raw Data)				
CC	$\bar{X}$	40.92	44.08	
n=13	s	17.78	19.38	
MAV	$\bar{X}$	45.13	46.73	
n=15	s	16.21	20.99	
TAV	$\bar{X}$	45.38	46.23	
n=13	s	20.98	22.67	
MAud	$\bar{X}$	45.08	42.23	
n=13	s	14.18	15.60	
Fmax		2.19	2.11	
F		0.20	0.15	
Durrell-Sullivan Reading Achievement: Word Meaning (Grade Level)				
CC	$\bar{X}$	3.95	4.13	0.18
n=13	s	1.19	1.27	0.17
MAV	$\bar{X}$	4.24	4.31	0.07
n=15	s	1.17	1.38	0.27
TAV	$\bar{X}$	4.12	4.25	0.14
n=13	s	1.21	1.22	1.01

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
MAud	$\bar{X}$	4.06	4.57	0.51
n=13	s	0.99	0.97	0.58
Fmax		1.48	2.02	34.33*
F		0.17	0.29	1.28

Durrell Sullivan Reading Achievement: Word Meaning (Raw Data)

CC	$\bar{X}$	25.38	28.00
n=13	s	13.10	14.60
MAV	$\bar{X}$	28.87	29.47
n=15	s	11.43	13.13
TAV	$\bar{X}$	27.31	29.15
n=13	s	12.05	13.59
MAud	$\bar{X}$	27.62	33.23
n=13	s	10.20	11.39
Fmax		1.65	1.64
F		0.21	0.38

Durrell-Sullivan Reading Achievement: Paragraph Meaning (Grade Level)

CC	$\bar{X}$	3.65	4.12	0.48
n=13	s	0.95	0.85	0.77
MAV	$\bar{X}$	3.75	4.28	0.53
n=15	s	0.82	1.38	0.91
TAV	$\bar{X}$	4.02	4.37	0.35
n=13	s	1.67	1.54	0.61
MAud	$\bar{X}$	3.94	4.35	0.42
n=13	s	0.74	0.79	0.59
Fmax		4.19	3.76	2.34
F		0.35	0.17	0.16

Durrell-Sullivan Reading Achievement: Paragraph Meaning (Raw Data)

CC	$\bar{X}$	14.77	16.08
n=13	s	6.90	5.63
MAV	$\bar{X}$	16.27	17.47
n=15	s	5.79	9.54

Source		Pre-Test	Post-Test	Gain
TAV	$\bar{X}$	17.92	17.92	
n=13	s	10.02	10.15	
MAud	$\bar{X}$	17.46	17.46	
n=13	s	4.68	5.54	
Fmax		4.58*	3.35	
F		0.52	0.12	
Durrell-Sullivan Reading Capacity: Total Score (Raw Data)				
CC	$\bar{X}$	74.85	82.64	7.79
n=14	s	11.31	12.06	5.89
MAV	$\bar{X}$	79.78	85.07	5.29
n=14	s	13.57	20.72	7.74
TAV	$\bar{X}$	69.69	76.08	6.39
n=13	s	15.81	13.39	14.09
MAud	$\bar{X}$	72.21	78.21	6.00
n=14	s	14.90	15.69	5.68
Fmax		1.48	2.95	6.15
F		1.18	0.91	0.20
Durrell-Sullivan Reading Capacity: Word Meaning (Raw Data)				
CC	$\bar{X}$	43.14	49.42	6.28
n=14	s	5.79	6.09	4.97
MAV	$\bar{X}$	46.35	48.21	1.86
n=14	s	11.01	11.14	2.98
TAV	$\bar{X}$	40.69	42.23	1.54
n=13	s	10.11	8.90	10.42
MAud	$\bar{X}$	41.57	44.42	2.85
n=14	s	7.90	8.73	4.19
Fmax		3.61	3.35	12.2*
F		1.08	1.90	0.96
Durrell-Sullivan Reading Capacity: Paragraph Meaning (Raw Data)				
CC	$\bar{X}$	31.71	33.21	1.50
n=14	s	6.71	7.70	3.16

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
MAV	$\bar{X}$	33.43	36.86	3.43
n=14	s	6.87	10.85	6.86
TAV	$\bar{X}$	30.54	32.31	1.77
n=13	s	5.09	5.40	5.86
MAud	$\bar{X}$	30.64	34.50	3.86
n=14	s	7.91	8.78	4.78
Fmax		2.41	4.03*	4.72*
F		0.55	0.59	0.12
Bond-Clymer-Hoyt:	Recognition Pattern (Raw Data)			
CC	$\bar{X}$	57.31	58.54	1.23
n=13	s	19.20	16.16	10.15
MAV	$\bar{X}$	58.07	62.71	4.64
n=14	s	18.55	16.87	10.46
TAV	$\bar{X}$	57.77	65.92	8.15
n=13	s	16.53	13.40	10.60
MAud	$\bar{X}$	63.14	67.43	4.29
n=14	s	15.64	10.15	7.42
Fmax		1.51	2.76	2.04
F		0.34	1.00	1.10
Bond-Clymer-Hoyt:	Visual Analysis (Raw Data)			
CC	$\bar{X}$	57.18	59.45	2.27
n=11	s	13.29	12.98	9.24
MAV	$\bar{X}$	60.08	63.00	2.92
n=13	s	12.01	11.30	6.55
TAV	$\bar{X}$	56.00	61.54	5.54
n=13	s	16.41	10.91	8.21
MAud	$\bar{X}$	54.00	58.36	4.36
n=11	s	15.17	14.58	11.18
Fmax		1.87	1.79	2.91
F		0.55	0.33	0.34

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
Bond-Clymer-Hoyt: Phonetic Knowledge (Raw Data)				
CC	$\bar{X}$	83.36	85.09	1.73
n=11	s	18.16	17.58	6.56
MAV	$\bar{X}$	69.85	82.30	12.45
n=13	s	21.70	22.28	13.83
TAV	$\bar{X}$	68.92	73.85	4.92
n=13	s	13.39	16.61	13.42
MAud	$\bar{X}$	76.45	80.00	3.55
n=11	s	10.87	14.58	8.18
Fmax		2.62	2.33	2.61
F		1.90	0.85	2.21
Bond-Clymer-Hoyt: Word Synthesis (Raw Data)				
CC	$\bar{X}$	12.73	12.09	0.64
n=11	s	4.92	5.84	4.34
MAV	$\bar{X}$	13.08	12.00	1.08
n=13	s	4.96	5.70	3.64
TAV	$\bar{X}$	11.61	12.38	0.77
n=13	s	5.01	5.49	4.19
MAud	$\bar{X}$	14.09	14.09	0.00
n=11	s	3.70	6.23	5.44
Fmax		1.83	1.28	2.24
F		0.62	0.34	0.45
PMA Total Score (Raw Data)				
CC	$\bar{X}$	88.55	104.00	15.4
n=9	s	28.61	24.17	10.03
MAV	$\bar{X}$	91.25	101.41	10.2
n=12	s	31.37	10.04	9.20
TAV	$\bar{X}$	83.90	95.30	11.4
n=10	s	18.01	15.02	10.74
MAud	$\bar{X}$	96.76	100.15	3.4
n=13	s	24.12	27.04	13.91
Fmax		3.04	7.29*	2.29
F		0.48	0.20	2.21

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
PMA Verbal Meaning (Raw Data)				
CC	$\bar{X}$	25.88	28.77	2.9
n=9	s	10.45	8.54	5.04
MAV	$\bar{X}$	29.75	31.17	1.4
n=12	s	9.24	11.46	5.79
TAV	$\bar{X}$	26.50	30.40	3.9
n=10	s	9.54	10.77	3.51
MAud	$\bar{X}$	27.23	27.00	0.23
n=13	s	7.42	8.19	6.98
Fmax		1.98	1.96	3.95
F		0.39	0.43	0.96
PMA Reasoning (Raw Data)				
CC	$\bar{X}$	26.33	28.33	2.0
n=9	s	8.00	5.39	6.28
MAV	$\bar{X}$	25.75	30.17	4.42
n=12	s	9.96	7.97	4.50
TAV	$\bar{X}$	22.40	27.70	5.3
n=10	s	6.28	3.68	5.53
MAud	$\bar{X}$	27.23	28.62	1.38
n=13	s	7.90	7.95	6.94
Fmax		2.52	4.68	2.38
F		0.70	0.27	1.13
PMA Perceptual Speed (Raw Data)				
CC	$\bar{X}$	9.00	13.44	4.4
n=9	s	6.96	6.44	4.92
MAV	$\bar{X}$	10.42	12.08	1.7
n=12	s	6.67	9.47	6.80
TAV	$\bar{X}$	9.4	9.0	0.4
n=10	s	4.44	5.97	3.76
MAud	$\bar{X}$	10.92	12.54	1.6
n=13	s	5.92	6.45	4.21
Fmax		2.47	2.16	3.28
F		0.24	1.38	1.53

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
PMA Spatial Relations (Raw Data)				
CC	$\bar{X}$	8.11	8.78	0.67
n=9	s	4.76	4.58	3.35
MAV	$\bar{X}$	7.92	8.25	0.33
n=12	s	5.18	4.03	2.57
TAV	$\bar{X}$	8.7	8.5	0.2
n=10	s	3.02	5.33	3.70
MAud	$\bar{X}$	12.08	11.54	0.54
n=13	s	4.42	4.63	3.27
Fmax		2.94	1.75	2.08
F		2.34	1.33	0.30
PMA Number Facility (Raw Data)				
CC	$\bar{X}$	19.22	25.11	5.9
n=9	s	9.67	11.16	5.11
MAV	$\bar{X}$	17.42	19.75	2.3
n=12	s	8.36	9.59	5.34
TAV	$\bar{X}$	17.10	19.70	2.6
n=10	s	8.46	9.25	4.52
MAud	$\bar{X}$	19.31	21.23	1.9
n=13	s	6.78	7.72	4.68
Fmax		2.04	2.09	1.40
F		0.22	0.71	1.33
"Speed" Portion of Pressey Diagnostic Reading Tests (Raw Data)				
CC	$\bar{X}$	15.77	20.08	4.31
n=13	s	7.72	10.69	5.78
MAV	$\bar{X}$	16.00	20.23	4.23
n=13	s	8.00	9.33	4.32
TAV	$\bar{X}$	17.36	21.27	3.91
n=11	s	12.89	15.51	3.01
MAud	$\bar{X}$	17.42	20.92	3.50
n=12	s	9.32	7.72	6.02
Fmax		2.78	4.04	3.99
F		0.01	0.03	0.07

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
IOTA Word Discrimination Test (Raw Data)				
CC	$\bar{X}$	33.10	33.40	0.30
n=10	s	10.24	8.64	3.17
MAV	$\bar{X}$	32.84	35.61	2.77
n=13	s	8.21	7.69	4.48
TAV	$\bar{X}$	33.58	35.00	1.42
n=12	s	8.97	8.34	4.52
MAud	$\bar{X}$	33.30	35.84	2.54
n=13	s	7.91	8.22	3.57
Fmax		1.67	1.18	2.04
F		0.02	0.20	0.90

Diagnostic Reading Test - Level 2: Total Score (Raw Data)

Source		Post-Test
CC	$\bar{X}$	39.92
n=12	s	13.62
MAV	$\bar{X}$	37.14
n=14	s	16.30
TAV	$\bar{X}$	40.54
n=13	s	12.56
MAud	$\bar{X}$	38.31
n=13	s	15.79
Fmax		1.68
F		0.15

Diagnostic Reading Test - Level 2: Word Attack (Raw Data)

Source		Post-Test
CC	$\bar{X}$	22.00
n=12	s	8.10
MAV	$\bar{X}$	17.35
n=14	s	3.25
TAV	$\bar{X}$	21.92
n=13	s	6.98
MAud	$\bar{X}$	20.08
n=13	s	9.00
Fmax		7.69*
F		0.84

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
Diagnostic Reading Test - Level 2: Comprehension (Raw Data)				
CC	$\bar{X}$		17.92	
n=12	s		6.57	
MAV	$\bar{X}$		17.64	
n=14	s		7.45	
TAV	$\bar{X}$		18.62	
n=13	s		7.37	
MAud	$\bar{X}$		18.23	
n=13	s		7.81	
Fmax			1.37	
F			0.01	
Diagnostic Reading Test - Level 3: Total Score (Raw Data)				
CC	$\bar{X}$		38.75	
n=12	s		15.64	
MAV	$\bar{X}$		41.17	
n=12	s		15.78	
TAV	$\bar{X}$		44.18	
n=11	s		11.28	
MAud	$\bar{X}$		44.77	
n=13	s		17.51	
Fmax			2.41	
F			0.04	
Diagnostic Reading Test - Level 3: Word Attack (Raw Data)				
CC	$\bar{X}$		25.25	
n=12	s		3.27	
MAV	$\bar{X}$		23.83	
n=12	s		9.04	
TAV	$\bar{X}$		28.91	
n=11	s		8.60	
MAud	$\bar{X}$		27.00	
n=13	s		13.94	
Fmax			18.15	
F			0.56	

APPENDIX B (continued)

Source		Pre-Test	Post-Test	Gain
Diagnostic Reading Test - Level 3: Comprehension (Raw Data)				
CC	$\bar{X}$		13.50	
n=12	s		7.19	
MAV	$\bar{X}$		17.33	
n=12	s		7.91	
TAV	$\bar{X}$		15.36	
n=11	s		4.28	
MAud	$\bar{X}$		17.77	
n=13	s		6.25	
Fmax			3.42	
F			0.09	

* Fmax significant at or beyond the 5% level

Appendix C

Distribution of Passage Difficulties by
Days of Presentation, Instructional
Phase, Experiment II

Day	Grade Level	Day	Grade Level
1	4	16	2
2	5	17	2
3	4	18	2
4	3	19	4
5	2	20	5
6	3	21	5
7	3	22	5*
8	3*	23	5
9	2*	24	5
10	3*	25	3
11	3*	26	2
12	4	27	2
13	3	28	5
14	4*	29	4
15	2	30	5

*Presented under normal/compressed conditions; remainder of passages presented under expansion conditions. Grades 2 and 3 passages were designated as "Easy," Grade 4 and 5 as "Hard."

Appendix D

Visual Survey Data

The use of bimodal (audio-visual) techniques in this study of course implied that the EMR subjects had the visual capability to read. The subjects were screened in this respect before inclusion in the study. However, the paucity of visual survey data for EMRs in the literature led to a decision to report these data for the subjects of the study as an appendix to the report.

A modified Ortho-rater* was used to administer a series of vision screening tests to experimental and non-experimental subjects for whom reading performance data were available. The experimental and non-experimental group data were separated inasmuch as the non-experimental group was defined as a group which read below Grade 2 level on the Recognition Pattern section of the Bond-Clymer-Hoyt Silent Reading Diagnostic Test (Form D-A); did not attain a SB, WISC, or WAIS IQ score between 54 and 86; wasnot between the chronological age of 12 years, 11 months and 17 years, 11 months; and had a major sensory defect or knowndegenerative neurological disease.

Comparative normal group data are reported.** The normal data (in press) are based on scores for 104 Grade 8 students with an average age of 13.8 and a mean IQ (Primary Mental Abilities Test) of 104. These data for Grade levels 1 through 8 indicated that most vision data scores were relatively constant, particularly for Grade 6 through 8. The Grade 8 values were chosen as standard since the mean age of this group most closely approximated that of the mentally retarded groups. The data for all groups are shown in Table D-1.

*Provided through the courtesy of Dr. Henry A. Knoll, Head of Research in Biophysics, Bausch & Lomb Optical Company.

**Through the courtesy of Dr. Helen M. Robinson of the University of Chicago.

Table D-1

Raw Score Means and Standard Deviations of Ortho-rater Scores for Normal Children and EMR Children Reading at or Beyond Grade 2 Level and EMR Children Reading Below Brade 2 Level

Visual Test	Group	N	Mean	s	No-Score
Right Eye Acuity (Far)	Normal	104	8.8	1.9	
	Reading	53	6.8	2.2	2
	Non-reading	27	6.3	2.6	3
Right Eye Acuity (Near)	Normal	104	9.9	1.5	
	Reading	54	7.5	2.4	1
	Non-reading	25	8.4	2.2	5
Left Eye Acuity (Far)	Normal	104	8.6	2.2	
	Reading	52	7.1	2.5	3
	Non-reading	25	6.2	2.3	5
Left Eye Acuity (Near)	Normal	104	10.3	1.7	
	Reading	52	8.1	2.5	2
	Non-reading	24	7.5	2.3	5
Lateral Phoria (Far)	Normal	104	7.4	1.9	
	Reading	52	7.3	3.4	3
	Non-reading	24	7.7	3.5	3
Lateral Phoria (Near)	Normal	104	5.5	2.1	
	Reading	52	6.3	3.3	3
	Non-reading	25	7.7	3.3	5
Accommodation/ Accommodative Convergence Ratio	Normal	104	8.7	1.5	
	Reading	52	5.0	1.1	3
	Non-reading	24	4.7	1.8	6
Stereopsis (Far)	Normal	104	4.2	3.0	
	Reading	35	3.3	2.7	20
	Non-reading	19	3.1	2.2	11

A number of subjects could not be tested because the range of the instruments was insufficient, or levels of vision were so low that the subjects could not make an adequate response, or suppression of vision in one eye was so pronounced that no measure could be obtained. The number of subjects who could not be scored is shown for each vision test for the retarded subjects. No comparable data are available in the normal data. Because of these anomalies, no statistical tests were attempted.

Visual Acuity

Visual acuity at both far and near was tested for each eye separately with a blank but illuminated target before the eye not being tested. In general, acuity for the normal group was higher than for the retarded groups. Standard deviations were approximately the same for all groups.

Lateral Phorias

Lateral phorias are measures of the extent to which the eyes deviate from a parallel position when the stimulus to fusion is removed. The mean laterial phorias at far distance were nearly the same for all groups. The retarded groups tended to show a wider range of variation, however.

The mean lateral phorias at near distance were similar for the normal and experimental groups. The near laterial phoris for the non-experimental groups was nearly two scale points higher than for the normal and experimental groups. The difference was in the direction of a tendency to greater divergence of the eyes at the reading distance when a stimulus to fusion was lacking. For clear binocular vision to obtain at the reading distance, a greater fusional convergence effort would be required. Standard deviations for the retarded groups were larger than for the normal groups.

Accommodation/Accommodative-Convergence Ratio

The amount by which the eyes converge per diopter of accommodation (A. C. A. ratio) was calculated from the near and far point lateral phoria measures. Mean scores for both retarded groups were similar and were approximately three prism diopters lower than for the normal group. The retarded groups required a greater amount of positive fusional convergence (ability to turn the eyes inward) in order to maintain clear binocular vision at the reading distance. The dispersion of scores for the experimental group was somewhat smaller than for the normal and non-experimental groups. The same relation for dispersion of scores held for the near point laterial phoria measures. The non-experimental group, however, showed a tendency to great divergence of the eyes relative to the point of fixation than either the normal or experimental groups.

Stereopsis

The stereopsis test measures the ability of a subject to judge distance from the small differences in the retinal images which are due to the separation of the two eyes. Approximately one-third of each retarded group could not respond to the easiest item on the test. Mean scores for the retarded groups were similar, and both were lower than for the normal groups.

The standard deviations of the stereopsis scores for both retarded groups were nearly as large as the means. The distributions were markedly skewed. One half of the subjects in each retarded group who were able to respond to the test achieved the lowest score on the scale. The lower visual acuity of the retarded groups probably influenced the mean stereopsis scores since relatively small letters must be discriminated in order to make a correct response. The large standard deviation in relation to the mean for the normal group suggests that the distribution of scores for that group was also skewed.

In summary, visual acuity, accommodation/accommodative ratio, and stereopsis scores were lower for the retarded than for the normal groups. Lateral phorias at the far point were similar for all groups with the retarded groups showing a greater dispersion of scores than the normal group. Lateral phorias at near point increased progressively in the divergent direction from the normal to experimental to non-experimental groups. In no case could the performance of retarded groups be said to be superior to the normal groups: the higher mean lateral phoria scores at both near and far points and the lower mean A. C. A. ratio scores place a greater than normal demand upon fusional convergence.

An additional indication that the visual skills of the non-experimental group were generally inferior to the experimental group is found in the frequency of no-scores for vision tests. While the non-experimental group was only half as large as the experimental group in seven or eight tests, the number of Ss who could not be tested in the non-experimental group was equal to or larger than the number of Ss who could not be tested in the experimental group. Conservatively, then, the incidence of no-test scores was at least twice as high as the non-experimental group, indicating a higher incidence of visual disabilities in this group. For the eighth test (stereopsis) the proportion of no-test scores was approximately one-third for both groups.

Implications for Reading

While far and near visual acuity for the retarded groups was, on the average, lower than for the normal group, these visual acuity impairments do not necessarily result in a lack of reading achievement. Persons with low visual acuity often compensate for the lack of clear vision by holding their reading closer to their eyes, thus obtaining a larger image. Such an adjustment increases the amount by which the two eyes must be turned inward in order to maintain single vision while using two eyes. However, the tendency for the eyes of the retarded groups to diverge at the near distance when the stimulus to fusion is eliminated suggests that, on the average, the retarded students must make a greater than normal effort to maintain single and clear vision at the reading distance.

While the expected finding is that there is some tendency for the eyes to diverge at the near distance, double vision is avoided where there is sufficient fusional convergence to overcome the tendency. Where fusional converge is inadequate, the coordination of the two eyes is lost and double vision occurs. Compensation for these difficulties may also be made. Double vision is prevented when the vision of one eye is suppressed, one eye is closed or covered, the head is turned so that the nose occludes one eye, or the reading is held so close that the double images are so far apart that one does not overlap the other. In some instances the effort to maintain single vision results in excessive stimulation to accommodation, so that print may be blurred, but not double.

The analysis of the accommodation/accommodative ratios for the three groups indicated that, on the average, less turning inward of the eyes was associated with focusing the eyes for the retarded group than the normal group. Thus, a looser linking of accommodation and convergence in the retarded places a greater demand upon positive fusional convergence if single vision is to be maintained. If, at the same time, lower acuity requires that the retarded student bring his reading closer to the eyes, an even greater demand is placed upon convergence. If these students do not have an adequate positive fusional reserve, discomfort, fatigue, and blurring or doubling of the print may occur, or the reader may become confused due to temporary doubling of the print which occurs as the eyes make a large saccadic movement to proceed from the end of one line of print to the beginning of the next line of print. Suppression of the vision in one eye may occur and binocular vision be "given up."

It is entirely possible that the potential benefits of bimodal stimulation in the present experiment were obscured or reduced by ocular coordination problems. Support for this hypothesis is found in the Robinson data (in press) which indicated that measurement of the positive fusional reserves, the lateral phorias, and the presence of far and near sightedness dintinguished achieving and non-achieving reading groups in a normal population. Further research involving measurement of the refraction of the eyes and the fusional reserves of the eyes is required in order to establish the relations between visual syndromes and achievement in reading.

APPENDIX E

A General Discussion of Issues Related to the Present Experiment

For better understanding the present research must be placed in perspective with respect to prior research and the problems of educating EMRs. Education, broadly defined, is the process of transmitting a culture from one generation to the next. The two basic determinants of an instructional program designed to transmit a given culture are:

(a) The goals of that culture. Such goals are operationally defined in terms of the behavior of those who are being and have been educated;

(b) The ability of the individual to manifest those behaviors under the appropriate conditions.

The present research is concerned with a subject population which, by definition, is deficient in the functional ability to display such behavior: The Educable Mental Retardate (EMR).

Heber (1961) uses the terms Mental Retardation (MR) in a generic sense and offers the following definition:

> "Mental retardation refers to a subaverage general intellectual functioning which originates during the developmental period and is associated with impairment in adaptive behavior." (p. 3)

Heber's definition is rather unique in that there is no assumption that there is an entailed relationship between measured intelligence and non-adaptive behavior. By implication, it precludes such an assumption in the absence of confirmatory evidence.

Heber (1961) offers the following caution in the use of his definition:

> "Within the framework of the present definition mental retardation is a term descriptive of the current status of the individual with respect to intellectual functioning and adaptive behavior. Consequently, an individual may meet the criteria of mental retardation at one time and not at another. A person may change status as a result of changes in social standards or conditions or as a result of changes in efficiency of intellectual functioning, with level of efficiency always being determined in relation to the behavioral standards and norms for the individual's chronological age group." (p. 4)

Sarason (1955) makes a distinction between Mental Retardation and Mental Deficiency:

> "(Mental retardation) refers to individuals who, for temporary or long stanging reasons, function intellectually below the average of their peer groups but whose social adequacy is not in question or, if it is in question, there is the likelihood that the individual can learn to function independently and adequately in the community.
>
> "(Mental deficiency) refers to individuals who are socially inadequate as a result of an intellectual defect which is a reflection of an impairment of the central nervous system which is essentially incurable." (pp 440-442)

Doll (1941) offers a more inclusive definition of mental deficiency:

> "The mentally deficient person is (1) socially incompetent, that is, socially inadequate and occupationally incompetent and unable to manage his own affairs; (2) mentally subnormal; (3) retarded intellectually from birth or early age; (4) retarded at maturity; (5) mentally deficient as a result of constitutional origin, through heredity or disease, and (6) essentially incurable." (p 214)

Heber's definition does not gainsay the existence of a population which meets Doll's criteria. On the other hand, conditions such as, incurability, constitutionality of origin, and retardation at maturity, cannot be descriptive of the current status of a child, or are themselves unknown. Thus it lacks operational clarity and, perforce, utility to the behavioral scientist. However, to the educator, the distinction between the mentally deficient and the mentally retarded must be made, even at the risk of its being invalid in an individual case. Instructional goals and techniques must vary as a function of a distinction (implicit or explicit) made between the two populations. In the case wherein there is confirmatory bio-medical evidence of central nervous system (CNS) pathology, the decision is facilitated (although not necessarily validated). In the case wherein such evidence is lacking, the special class teacher, lacking in training in either biological or psychological diagnosis, is often by default the sole determinant of the distinction.

The distinction between the definitions of MR offered by Heber and by Sarason are as follows:

(1) Heber's definition includes the mentally deficient.

(2) With respect to the social adequacy of the adult:

(a) Sarason's definition excludes those who are inadequate.

(b) Heber, however, includes social inadequacy as a necessary condition.

(3) With respect to the social adequacy of the child:

(a) Sarason's definition, again, precludes defining the socially inadequate as retarded.

(b) Heber's position is that social adequacy may, but need not, be a relevant criterion. If it is relevant, it is considered as definitional attribute of the MR.

The conflict is more apparent than real as "social adequacy" is the weakest link in the operational chain. The issue is still of import to the teacher, however. What he teaches and how he teaches will vary according to whether or not the child is deemed mentally retarded.

The focus of the distinction in this case, however, is between the MR and the "normal" rather than between the mentally retarded and the mentally defective.

In operational terms, the EMR refers to a population of school aged mental retardates:

(a) Whose IQ falls between 55 and 85, and

(b) Whose impairment in learning and adaptive behavior is such that placement in a special class has been deemed necessary.

Heber's defintion of MR can be accepted as it admits of, but does not entail, etiological distinction of dubious scientific merit. The EMR's behavioral inadequacies are not so severe as to imply constitutional defect nor to preclude remediation.

It is assumed that the EMR can manifest some, but not all, of the behaviors which, on the basis of consensual validation and normative standards, are defined as acceptable and/or desirable.

A child who is classified as an EMR has the potential to develop the minimal skills needed to operate as an adaptive member of our society, but, because of his limited ability to profit from experience, special methods of teaching are needed.

Because of the need for more intensive instruction, it is unlikely that all of the potentially desirable behaviors of the EMR can be taught during the developmental period. It is probably necessary to restrict the teaching to the most necessary and functional behaviors. While it is suggested that special instruction is needed and it ought to be restricted to vital needs, it is not meant to imply that the functional ability (as measured by the IQ) is an adequate reflection of innate, biological potential.

The rejection of the validity (scientific utility) of an assumed monotonic relationship between measured functional ability and innate capacity (see for e.g., Masland, Sarason, and Gladwin, 1958; Hunt, 1961; Liverent, 1960) presents a crucial issue to those who would provide curricula for the EMR. The issue is: To what extent does the restricted instructional program employed in teaching the EMR preclude the development of a latent potential for "normal behavior" which was obscured by inadequate environmental stimulation? At present, there is no adequate answer to this problem. However, one source of relevant data is the research literature related to the efficacy of special classes for the EMR.

Johnson (1962) has reviewed the literature on the efficacy of special class placement for the EMR and has raised some disturbing issues relative to present instructional goals and procedures used in special education.

A considerable amount of professional effort has been expended upon the establishment of special classes for the EMR, attempts to justify such segregation and attempts to develop special curricula for such classes which take into consideration the "characteristics and limitations" of the EMR. A considerable body of research on the efficacy of special classes suggests that whatever support can be garnered in their favor is derived from rather equivocal evidence that EMRs in special classes benefit in the area of social and personal development. Serious measurement problems, however, vitiate the significance of such findings. There is no basis in the research literature to justify special classes when achievement in school subject material is taken as the criterion. There are considerable data which suggest that there is no significant difference, in many cases, between the achievement of EMRs place in special classes and those retained in the regular grades.

When significant differences have obtained they have clearly favored the EMR retained in the regular grades (see also, Kirk, 1964).

Proponents of special classes have validly contended that the efficacy studies have been replete with problems such as: criterion and measurement index selection; sampling bias; teacher adequacy; nature of the instructional program; and a host of other relevant issues. However, taking the findings at face value, what are some of the factors which might account for them:

Brown and Lind (1931) found that low IQ (50-59) children achieve above MA expectancy, children with IQs ranging from 60 to 69 achieve near expectancy, while those with IQs from 70 to 79 fell far below what was expected of them. Brown and Lind suggest that instruction is directed at the average performance level of the class. Thurstone (1959) also found that lower IQ children profited more from placement in special class. Thus, a lower level of instruction and less opportunity to profit incidentally from contact with brighter CA peers in the special class may account for the differences which have been found to favor regular class placement.

Special class teachers are taught to adapt curricular goals and content in a way that may well make the efficacy study criteria inappropriate. In addition, they are also taught to adapt instruction to MA expectancy. A number of studies (e.g., Foss, 1938; Durrel, 1940; Dodge, 1940; Gilman, 1941; Speidel, 1958; and Biggins, 1961) suggest that using the MA as an index of expectancy results in too low an instructional level as all of these studies suggest that grade level attainment in listening comprehension (on the Durrell-Sullivan Reading Capacity Test) exceeds MA expectancy.

It may be that teachers, in selecting or developing reading material appropriate for the independent reading level of the EMR often unwittingly lower the instructional level well below even the conservative MA expectancy level. Thus, we in special education may tend to preclude the maximal development of special class EMRs.

Research on Reading

Reading has been a primary focus to attention of research related to EMRs (Kirk, 1965). On the basis of an exhaustive review of the literature, Dunn (1956) concluded that EMRs tend to read below MA expectancy but that there is also evidence that, given special attention, the reading achievement of the EMR should be expected to equal or exceed such expectancy. It should be again noted in passing that taking the MA rather than the CA as a criterion for expectancy minimizes the discrepancy between expectancy and achievement. Kirk (1964) reports that studies by Braem (1931), Hegge (1934), McIntyre (1937), and French (1950) have demonstrated that a phonic approach to remedial reading instruction is effective with EMRs. Coleman (1938) and Storey (1936) report satisfaction with an eclectic technique which stressed word recognition. Kirk (1933) found that EMRs profited from a kinesthetic method (over a sight approach) on retention of word recognition, but that there was no difference in initial learning.

Dunn (1956) compared the performance of a group of EMR boys with a group of normals of equal MA on various reading tests. He found the normals superior to the EMRs on the following overall reading achievement: faulty vowels; omissions of sounds; words aided and refused; context clues; tachistoscopic presentation of words and phrases; auditory acuity; visual acuity; social and emotional adjustment; economic status; social class; and emotional climate of the home. Vance (1956) and Capobianco and Miller (1958) matched on CA, MA, and IQ brain-injured and non-brain-injured MRs and found no significant differences in reading, type of errors, auditory acuity, or visual perception.

Thus, Kirk (1964) has concluded that we cannot assume measured organic deficit to underly the EMR's reading disability and that studies such as that of Frey (1961) suggest that systematic education can compensate for deficits which are assumed to be a function of organic damage. It would appear that research literature on the reading achievement of the EMR supports the following general conclusions:

(a) In general, EMRs do not perform up to minimal estimates of MA expectancy.

(b) There is no valid basis for assuming that the EMR cannot learn to read up to MA expectancy.

(c) No one has adequately demonstrated a superiority of a given teaching technique with this population.

(d) In light of present knowledge, a priori assumptions as to any irremedial organic limitations on the EMR's ability to learn to read are, at best, highly tenuous.

One point which may well be unique with this population is worthy of mention. The contention that EMRs can learn to read does not, in the mind of many special educators, serve to validate learning to read as a goal of the school program for the EMR. Specification of "learning to read" as a goal generally entails the tacit assumption that eventually the child will "read to learn." How well EMRs can "read to learn" is a moot point. Comprehension is generally the weakest part of the EMR's reading skills profile. A prognosis of 4th grade competency is generally taken as the minimal standard of educability. Many EMRs who do attain 4th grade level overall on a standardized achievement test have a reading comprehension level of 1st or 2nd grade or less. Are such children "really" educable? Is the effort entailed in teaching the EMR to read spent wisely if there is little surety that reading will become a functional tool for learning?

The above issue is a vital one. The most direct attack on the problem would entail specification of the nature of the intellectual and psycholinguistic processes of the EMR. However, the "answer" is not apt to be forthcoming in the near future in terms of giving any sense of direction to service personnel. Another approach, however, is possible. If an efficient method of teaching reading in a short period of time can be developed and validated <u>or</u> if reading skills can be enhanced as a sort of by-product of attempts to develop functional content knowledge and social and vocational skills, perhaps the issue can be divested of any great import and may even serve to stimulate efforts directed toward teaching the retardate to "read to learn."

The present research describes an attempt to enhance growth in reading skills as an ancillary function of the development of functional content knowledge by employing bimodal stimulation of redundant material. Our instructional focus was upon the comprehension of the content of the lessons. There was no attempt to teach word attack skills. The goal was to stress "reading (and listening) to learn" rather than "learning to read." This distinction breaks down with respect to the introduction of unfamiliar vocabulary. In the course of the lessons, the children were taught the meaning of words they were not apt to know in order to permit comprehension. It was necessary to insure that the children could recognize a word prior to explaining its meaning. As "word recognition" is a reading skill, the functional dichotomy breaks down. To the extent that such word recognition was entailed in the teaching of the lessons, it is clear that teaching was addressed to the acquisition of a "skill" as well as the informational content of the passage.

To summarize, the following observations are offered:

(1) The EMR has generalized behavioral inadequacies which are intimately associated with language dysfunction.

(2) The inadequacies are relative to the demands of society.

(3) There is no valid basis for assuming that the inadequacies cannot be overcome; thus, it is incumbent upon the special educator to provide a more suitable educational environment.

(4) On the other hand, we have not been able to demonstrate that remediation is possible by environmental manipulation.

(5) Any attempt at remediation will make demands upon the child. Since the impact of such demands are unknown we must be very careful with respect to (a) the nature of the demands, (b) the means of enforcing the demands, and (c) the way in which we interpret the child's response to the demands (especially with regard to inferring negative motivation or an inherent lack of ability). We must give serious consideration to the possibility that we have not developed appropriate teaching techniques.

(6) There is considerable evidence to suggest that the retardate can "learn to read." While there is little evidence that he "reads to learn," this does not necessarily invalidate learning to read as an educational goal.

(7) Since the functional utility to the EMR of reading as a learning tool is open to question, the special educator must be sure to (a) establish a hierarchy of educational goals in terms of their functional utility, and (b) insure that the teaching of reading does not displace educational goals which are of equal or greater value.

(8) The determination of utility is an empirical question which cannot be answered without attempting to teach the EMR to read. While this validates learning to read as an educational goal, it sets a limit in that unless progress in "reading to learn" is achieved, further attempts to enhance reading skills should not be pursued since they may lead to goal displacement.

(9) In order to avoid goal displacement, the special educator must: (a) develop efficient means, in terms of instructional time and

effort, of teaching the child to "learn to read" (i.e., encode graphemes); (b) evaluate the impact of her teaching on pupil performance, noting particularly the relationship between the child's word attack skills and his listening and reading comprehension. As the development of specific reading skills approach an asymptote, there should be a monotonic increase in reading comprehension; (c) if an increase in comprehension obtains, the child can read to learn and the validity of learning to read as an educational goal is not in question; but (d) if an increase in reading comprehension relative to word attack skill acquisition and listening comprehension does not obtain, further instruction in reading skills is not apt to be functional. At this point, "learning to read" must be dismissed as a valid instructional goal or at least be deemed ancillary to the development of skills and acquisition of content needed for social and vocational integration into society. Thus, subsequent instruction involving "reading" cannot involve reading as an instructional medium for the acquisition of appropriate content nor be of such a character that the acquisition of reading skills are viewed as a sufficient educational goal of that instruction.

(10) The present experiment was an attempt to determine if bimodal (audio-visual) stimulation with redundant information which was within the listening comprehension level but above the reading comprehension level of the EMR would (a) facilitate comprehension of the content of the material; and/or (b) yield an increase in reading skills as an ancillary by-product. Of primary interest was an assessment of performance by EMRs as a function of bimodal stimulation wherein the auditory materials are taped and presented by means of an apparatus which can vary the rate of presentation without pitch distortion. The performance was evaluated relative to that of three contrast groups: (a) a group exposed to bimodal stimulation wherein the auditory material was presented by the teacher; (b) a group exposed to auditory stimulation only; and (c) a classroom control group.

Prior Research on Audio-Visual Stimulation

Hartman (1961b) and Travers, et al. (1964) have published rather extensive reviews of research and theory related to audio-visual communication. Hartman offered the following comment as to the comparative effectiveness of audio versus visual (print) stimulation:

> "Audio is a more effective channel than print when the information presented is simple and easily understood by the subjects, and for illiterates and semi-illiterates (e.g., children) regardless of the difficulty of the information. Print shows increasing advantage over audio for literate subjects roughly proportional to the increasing difficulty in their comprehension of the material." (pp. 239-240)

These conclusions are based upon the results of over 50 studies involving the use of nonsense syllables and digits, meaningful words, and connected discourse. Hartman suggests that the "referability of visual displays is an asset especially as the difficulty of material increases."

With respect to the efficacy of bimodal (audio and print) versus unimodal, audio (A) or print (V), stimulation with redundant information, Hartman (1961b) reported that of nine studies comparing AV versus A channel presentation:

Four studies favored AV presentation.

Two studies favored A presentation.

Three studies reflected no differences.

Hartman (1961a) presented evidence that bimodal (audio-print) stimulation with redundant material yields superior performance. Hartman contended that the general superiority of bimodal exposure obtains in spite of the fact that testing has generally been unimodal (primarily print). This fact suggests that Hartman's (1961b) stimulus generalization interpretation of the facilitation associated with bimodal stimulation is not adequate to account for the data. A rigorous stimulus generalization model would not predict superior performance under condition of bimodal stimulation unless bimodal stimulation were employed in the test situation.

Travers, et al. (1964) do not share Hartman's (1961a, b) conclusion with respect to facilitation as a function of bimodal stimulation with redundant material. They contended that a host of methodological inadequacies vitiate the significance of most of the studies in this area.

Travers, et al. (1964) reported that:

> "In the review of the early studies in this area by Beach and Day (1950) several factors were designated as important in designing an experiment in this area. These factors were: (a) the meaningfulness of the stimulus material--the more meaningful material favors the audio channel; (b) I. Q. --a higher I. Q. favors the visual channel; (c) reading ability--a high reading ability favors the visual channel; (d) age--as age increases the relative effectiveness of the visual channel increases; (e) the method of testing used to determine the amount of learning that has gone on--visual testing favors the visual channel, audio testing favors the audio channel, etc.; (f) the interval between the learning task and the testing--a shorter period favors the visual channel, whereas a longer period favors the audio channel; (g) referability--visual displays have greater referability than auditory inputs; this might account for some of the advantage of the visual channel in some of the earlier studies; (h) organized material (redundant) vs. non-organized material favors the audio channel; (i) and the criteria used to determine when the task has been completed--immediate memory favors the visual channel, and the amount retained over time favors the audio mode." (pp. 617-618)

Travers and his associates go on to report two studies conducted at the University of Utah which were designed to avoid the procedural shortcomings of prior studies. The results of these studies suggest no superiority associated with AV stimulation. In fact, under conditions of short stimulus exposure duration, there was some evidence of inferior performance. In addition, there was some evidence that audio materials are sometimes ambiguous as a function of inadequate phoneme discrimination and the lack of an isomorphic relationship in phoneme-grapheme correspondence.

None of the studies by Hartman (1961b) or Travers, et al. (1964) have been discussed in detail as their salience to the present work is limited by two concerns: (a) these studies have generally not involved the bimodal presentation of connected discourse; and (b) they have not involved the population of interest, the EMR.

Empirical evidence to support Hartman's conclusions is, at present, inadequate, but the possibility of such a finding warrants exploration in both empirical and theoretical terms. Two possible interpretations of the enhanced performance associated with bimodal presentation and unimodal testing may be suggested:

(1) There may be a summation of cues or traces involved in bimodal stimulation which are associative and central in character and which render the test stimulus and response modalities irrelevant, particularly in the case of connected discourse.

(2) It may be that, as a function of the interaction of historical and/or ahistorical variables, there exists in a given experiment population Ss who have a differential opportunity to acquire information and of test response modality proficiency rather than a summation of information from both modalities.

In general, an empirical test of these two alternatives is extremely difficult to achieve since attention is a response-defined variable. It would be necessary to induce differential responses to the stimuli presented via the two channels in order to determine differential attention. As there is no reason to assume that attentional processes are constant across conditions of presentation of redundant versus competing information, research involving response competition as a basis for inferring differential attention cannot serve to resolve the issue.

The Role of Effectiveness and Efficiency in the Selection of Communication Modalities

If the efficiency of communication is construed as a paramount goal, among the relevant factors are:

(1) The relative cost of equipment and material;

(2) The availability of transmission modes to a target audience;

(3) The size and geographic dispersement of that audience;

(4) The relative command the target audience has over the graphemic versus the phonemic code;

(5) The propensity of the target audience to search for the information one wishes to communicate;

(6) The range of rate of presentation over which effective communication is possible;

(7) The distribution and amount of time the target audience will remain responsive to the transmission of information;

(8) The behavioral response anticipated or desired of the target audience;

(9) The source's need for feedback in order to determine subsequent transmission.

The above list is extensive but hardly exhaustive. Of course, many of these factors are related to the effectiveness as well as the efficiency of communication. But, in general, this relationship can be summed up as follows: In order to be efficient, the information process must be effective. In this context, effectiveness must be cast in terms of increments attained above the minimal criterion and the cost, in terms of time, material, and effort, needed to reach or go beyond the minimal goal. In the present case, our minimal goals must be defined in terms of the present effectiveness and efficiency of communication in the special class.

The Effectiveness and Efficiency of Communication in the Special Class

Our target population is the EMR. Our communication efforts are directed toward the "education" of the EMR. Education, as a societal institution and function, is necessary only to the extent that the present and posited future behavior of the individual must be oriented and directed toward socially valued goals. These goals, and the options and sanctions the educator must have in order to obtain these goals, vary drastically with the complexity of the culture.

The "relative inadequacy" of the EMR is in terms of minimal behavior requirements entailed in adapting to a complex culture.

In the interest of effective education, we must insure the development of the ability to use and respond to speech - the phonemic code of our language. In the interest of efficient education, it is useful to develop the ability to read and write - to use and respond to the graphemic code of our language. With respect to both the goals of our culture and the discrepant functional abilities of the EMR, it is clear that the development of the ability to effectively encode and decode speech must be insured before the relative efficiency of developing reading as a functional tool can be accepted as a valid educational goal.

One of the principal advantages inherent in the segregated special class is a gain in psychological distance from demands cast in terms of what the child should do as defined in terms of MA expectancy or the normative standards implicit in the modal behaviors of the EMR's CA peers. As a consequence, both teacher and child gain some latitude. If the teacher is at all observant, intimate daily intercourse with the child offers some insight into what the child can and does do rather than being influenced primarily by what the child "should be able to do."

However, teachers, as a class, have been taught:

(1) to think in terms of what the EMR should be able to do and to define goals in terms of MA expectancy; and

(2) to value reading as an end unto itself.

The perceived discrepancy between MA expectancy and reading comprehension, by definition, is a function of educational retardation rather than mental retardation and is presumed to be remedial. The significance of the child's "overachievement" in listening comprehension is not evaluated, and intensive effort is directed at teaching the child to "learn to read." The teacher cannot teach "reading comprehension" per se, so he focuses upon word attack skills. These skills are brought more in line with "expectancy." Still, the child does not "comprehend" what he reads. So, the instructional content level is lowered to insure comprehension. The child begins to complain on two counts: (a) he "knows" he can't learn to read from past experience and (b) school consists of a lot of baby stuff that has no functional relationship to "the real world." The teacher faces the task of tieing non-functioning skills to functional goals defined in terms of future social roles and vocational pursuits. There is little intrinsic motivation tied to school work. Teacher attempts to motivate the child by saying, "You have to do this if you want to get a job!" Now the child knows it is important to get a job so he goes along with the teacher. He doesn't know why it's important; but he knows it must be important because everybody agrees with the teacher that it is important.

The problem that the educator faces in transmitting our culture is that our dynamic culture won't stand still long enough to be transmitted. In the case of the EMR, this issue is of vital import as it relates to the validity of our classifying the EMR as educable.

The purpose of this discussion is to point out the fact that the exploration of basic empirical and theoretical issues related to audio-visual communication transcends the level of being a purely academic concern. It is imperative that we develop effective and efficient means of communication with the EMR. The social competence of a significant number of our fellow human beings is at stake.

We must take, as a point of departure, the position that, at present, communication in the special class is seldom effective and hardly even efficient. Inadequate as such communication may be, it must serve as a baseline with which to evaluate the utility of alternative communication media. There may well be a discrepancy between the relative effectiveness and efficiency of communication media employed for various purposes. With respect to the EMR population, our primary focus must be on the development of effective techniques before the question of efficiency can be validly posed or adequately resolved.

References

Andrus, Ruth M. Diagnostic study of reading abilities in special classes and suggestions for improvement of difficulties. Unpublished masters service paper, Boston Univer., 1942.

Biggins, Mildred E. A comparison of listening comprehension and reading comprehension in second and third grades. Teach. Col. J., 1961, 33, 54, 55.

Braem, Helen R. An experiment in the methodology of teaching reading to mental defectives of moron and borderline level. Proc. Am. Assn. Study of Feebleminded, 1931, 36, 92-99.

Brown, A. W. and Lind, C. Scholastic achievement in relation to mental age: A comparative study. J. educ. Psychol., 1931, 22, 561-576.

Burt, Cyril. The backward child. New York: D. Appleton-Century, 1937.

Capobianco, R. J. and Miller, D. Y. Quantitative and qualitative analyses of exogenous and endogenous children in some reading processes. USOE, Coop. Research Program, Syracuse Univer., 1958.

Chipman, Catherine E. The vocabulary of mental defectives. Proc. Am. Assn. ment. Def., 1935, 40, 485-504.

Coleman, Marion U. Remedial reading for special groups. Proc. Am. Assn. ment. Def., 1938, 43, 123-127.

Dodge, Harriet E. A comparison of the Stanford-Binet mental age and hearing comprehension ability of dull children. Unpublished masters thesis, Boston Univer., 1940.

Doll, Edgar A. The essentials of an inclusive concept of mental deficiency. Amer. J. ment. Defic., 1941, 46, 214.

Dunn, L. M. A comparison of the reading processes of mentally retarded and normal boys of the same mental age. Monogr. Soc. Res. in Child Develpm., 1956, 19, Lafayette, Indiana: Child Dev. Pub.

Durrell, Donald D. Do dull children have high accomplishment quotients? Official Report of 1940 Meeting, Amer. Educ., Washington, D. C., 1940, 58-66.

Edwards, A. Experimental design in psychological research (Rev. ed.) New York: Holt, Rinehart and Winston, 1965.

Foss, Gertrude M. Language comprehension skills of mentally retarded children. Unpublished masters theses, Boston Univer., 1938.

French, E. L. Reading disability and mental deficiency: A preliminary report. Training Sch. Bull., 1950, 47, 47-57.

Frey, R. M. Reading behavior of brain-injured and non-brain-injured children of average and retarded mental development. Unpublished doctoral dissertation, Univer. of Illinois, 1961.

Gilman, A. Mental test variations among dull children. Unpublished masters thesis, Boston Univer., 1941.

Harris, A. J. How to increase reading ability (Fourth ed. Rev.) New York: McKay, 1961.

Hartman, F. R. Single and multiple channel communication: A review of research and a proposed model. Audio Visual Commun. Rev., 1961, 9, 235-262.

Heber, Rick. (Ed.) A manual on terminology and classification in mental retardation (2nd ed.) Monog. Suppl. Amer. J. ment. Defic., 1961.

Heckelman, R. G. A neurological impress method of reading instruction. Merced County Schools, Calif. (March, 1964). A report to the Annual Workshop on Reading Research conducted by the Committee on Diagnostic Reading Tests.

Hegge, T. G. Special reading disability with particular reference to the mentally deficient. Proc. Amer. Assn. ment. Defic., 1934, 39 297-343.

Hill, Iverna. A survey of reading difficulties of mentally retarded children in Schenectady, New York. Unpublished masters thesis, New York State Coll. for Teachers, Albany, 1939.

Hoyt, Mildred. Mental age and school attainment of 1007 retarded children in Massachusetts. J. educ. Psychol., 1924, 15, 297-301.

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

Johnson, G. O. Special education for the mentally handicapped: A paradox. Except. Child., 1962, 29, 62-69.

Kelly, Elizabeth M. Comparative study of achievement in reading of 100 normal and 100 subnormal children. Unpublished masters thesis, New York Univer., 1934.

Kendall, M. G. Rank correlation methods. London: Griffin, 1948.

Kirk, S. A. The influence of manual tracing in the learning of single words in the case of subnormal boys. J. educ. Psychol, 1933, 24, 525-535.

Kirk, S. A. The effects of remedial reading on the educ. progress and personality adjustment of high-grade mentally deficient problem children: Ten case studies. J. of Juv. Res., 1934, 18, 140-162.

Kirk, S. A. Research in education, In Heber, R. and Stevens, H. A. (Eds.) Mental retardation: A review of research. Univer. of Chicago Press, 1964.

Kirk, Winifred D. and Kirk, S. A. The influence of the teacher's handedness on children's reversal tendencies in writing. Pediatric Seminar and J. Genet. Psychol., 1935, 47, 473-477.

Lindquist, G. F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin, 1953.

Liverent, S. Intelligence: A concept in need of re-examination. J. counsel. Psychol., 1960, 24, 101-110.

McIntyre, E. M. Teaching of reading to mentally defective children. Proc. Amer. Assn. ment. Defic., 1937, 42, 56-67.

Masland, R. L., Sarason, S. B., Gladwin, T. Mental subnormality, New York: Basic Books, 1958.

Merrill, Maud A. The ability of the special class children in the 3 Rs. Pediatric Seminar, 1918, 25, 88-96.

Mintz, Alexander. Reading reversals and lateral preferences in a group of intellectually subnormal boys. J. educ. Psychol., 1946, 37, 487-501.

Monroe, Marion. Children who cannot read. Univer. of Chicago Press, 1932.

Mullen, Frances A. The reading ability of the older, ungraded pupil. Chicago: Bur. Mentally Handicapped Children, Chicago Public Schools, 1952.

Renshaw, Samuel. The abilities of pupils in Detroit pre-vocational classes. J. educ. Psychol., 1919, 10, 83-94.

Sarason, S. B. In W. W. Cruickshank (Ed.) Psychology of exceptional children. New York: Prentice Hall, 1955, 440-442.

Scarborough, W. H. The incident of reading retardation among 1182 mentally handicapped children. Paper read at Amer. Assn. ment. Defic. Fall, 1951.

Siegel, S. Non parametric statistics for the behaviorial sciences. New York: McGraw-Hill, 1956.

Speidel, Elizabeth. Language achievements of mentally retarded children. Unpublished doctoral dissertation, Boston Univer., 1958.

Spicker, H. Listening comprehension and retention of intellectually normal and retarded children as a function of speaking rate and passage difficulty. Unpublished doctoral dissertation. Nashville, Tenn: George Peabody College for Teachers, 1963.

Storey, Bernice L. Analytical appraisal of a remedial program in reading for pupils with an IQ of 90 or below. Unpublished doctoral dissertation, Univer. of Pittsburgh, 1936.

Thurstone, Thelma G. An evaluation of educating mentally handicapped children in special classes and in regular grades. Chapel Hill: Univer. of North Carolina, 1959.

Travers, R. M. W. (Ed.) Research and theory related to audio visual information transmission (Interim Report) USOE Contract No. 3-20-003. Univer. of Utah, 1959.

Vance, H. S. Psychological and educational study of brain-injured and non-brain-injured mentally retarded children. Unpublished doctoral dissertation, Ohio State Univer., 1956.

Walker, Helen, and Lev, J. Statistical inference. New York: Holt, Rinehart, and Winston, 1953.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Witty, P. A. and McGafferty, Estelle. Attainment by feebleminded children. Education, 1930, 50, 588-597.