The hypothesis/test procedure training materials in reading and word recognition were field tested with 40 educable mentally handicapped children (9 to 13-years-old) and 29 control Ss. Three classroom teachers used the materials for approximately one-half hour a day for 15 weeks. Information was collected through weekly conferences, observations of the classes, and pre- and posttests of skills. Data showed that the students using the experimental materials had acquired a better set of competent skills comprising the hypothesis/test word recognition process than the students not using the materials. Results of the component skills test did not prove unequivocally that the materials and instructional program produced higher gains on the subtests by the classrooms using the materials. Teacher conferences and classroom observations indicated that it is necessary to assess the impact of the hypothesis/test materials with readers not so advanced in the subskills of these materials, and that more cloze stories should be incorporated into the later units. (GW)
INITIAL FIELD TEST AND FEASIBILITY STUDY OF THE HYPOTHESIS/TEST WORD RECOGNITION PROCEDURES IN THE SPECIAL EDUCATION CLASSROOM

Christine Mueller and S. Jay Samuels
University of Minnesota
Research, Development and Demonstration Center in Education of Handicapped Children
Minneapolis, Minnesota

December, 1974

The research reported herein was performed pursuant to a grant from the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education, and Welfare to the Center for Research, Development and Demonstration in Education of Handicapped Children, Department of Psychoeducational Studies, University of Minnesota. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.
The University of Minnesota Research, Development and Demonstration Center in Education of Handicapped Children has been established to concentrate on intervention strategies and materials which develop and improve language and communication skills in young handicapped children.

The long term objective of the Center is to improve the language and communication abilities of handicapped children by means of identification of linguistically and potentially linguistically handicapped children, development and evaluation of intervention strategies with young handicapped children and dissemination of findings and products of benefit to young handicapped children.
Acknowledgements

The investigators would like to express their gratitude to the Minneapolis Public Schools personnel, especially, Dr. Richard Faungt—Director of Research, Dr. Jerry Gross—Assistant Director of Special Education, Mrs. Mae Peterson—Assistant Coordinator for Special Classes, Ms. Nita Grismer—Coordinator for Resource Programs and Special Classes, Mr. Sidney Buchanan—Principal of Standish School, Mrs. Chlores Johnson—Staff Member of Standish School, Mr. Donald Ravell—Principal of Loring School, Mrs. Dorothy Hegdal—Staff Member of Loring School, Ms. Janice Anderson—Principal of Wenonah School, Ms. Barbara Liszt—Staff Member of Wenonah School, Mr. George McDonough—Principal of Bethune School, Mrs. Daphne Ekstrom—Staff Member of Bethune School, Ms. Margaret Lincoln—Principal of Clinton School, Ms. Arlene Frieler—Staff Member of Clinton School, Mr. Ted Pollard—Principal of Harrison School, and Mrs. Bertha Smith—Staff Member at Harrison School. The investigators would also like to express their appreciation to the participating children whose cooperation made the present study possible.

Thanks also to Teara Archamety and Anne Morrow for their suggestions during the writing of the Hypothesis/Test materials.
A gap exists between educational research which has implications for the classroom and the actual implementation of these experimental results in classroom curricula. Curriculum developers attempt to bridge this distance between the educational researcher and the classroom teacher. The tasks involved include writing classroom materials based on experimental research implications, field testing these materials in an actual classroom, revising the materials, and field testing at least once again. This sequence of events, under the direction and supervision of curriculum developers, is necessary before materials are ready to be published and packaged for dissemination. After materials are once published and disseminated, the use of the materials as the publisher intended is yet another area of concern for the curriculum developer if he wishes to actually influence practices in the classroom.

Hess (1974, p. 10) states, "Surely one necessary element for effective change in our schools is good curricula that has been carefully tested, revised, and is usable." This paper describes the initial field testing of a partial reading program in Educable Mentally Retarded classrooms. The sequence of the development of the materials to this point and the proposed further development can be described as
The project described in this paper fits into the sequence of curriculum development at step 3.

Several years of experimental study at the University of Minnesota's Research, Development and Demonstration Center has shown that the Hypothesis/Test Model of word recognition can be broken down into behaviors that can be taught to EMR (Educable Mentally Retarded) children. (The description and rationale of the Hypothesis/Test Model of word recognition follows in the next section.) Archwamety and Samuels (1973) worked with retarded subjects using a randomized experimental and control groups design. These mentally retarded experimental subjects were taught to use the word recognition strategy derived from the Hypothesis/Test Model of word recognition. The results showed that experimental subjects demonstrated significantly better word recognition subskills than the controls.

Dahl, Samuels and Archwamety (1973) used normal subjects who were the poorest readers in their school to further test the effectiveness of the Hypothesis/Test Model strategy to levels of automaticity. (See the next section for the relation of automaticity to reading comprehension.) In this study it was assumed that normals and retardates must go through highly similar processes in reading acquisition, and that the subskills
required for acquisition would be similar in both cases, but that
retarded subjects would probably take longer in acquiring these
skills.

The results (Dahl, Samuels and Archwamety, 1973) indicated
that children trained in Hypothesis/Test procedures were superior
in a number of ways. First of all, on tachistoscopic word recogni-
tion which is considered to be a sensitive measure of speed of
processing words, it was found that the experimental group was
superior to its control on this variable.

Another important variable was comprehension. Bormuth (1966)
has evidence indicating that the cloze technique is in many ways a
superior measure of comprehension in comparison to traditional
comprehension measures. In this study the Hypothesis/Test-trained
students were superior on this measure as well.

The procedures for training children to use Hypothesis/Test
strategies are straightforward, simple and inexpensive. These pro-
cedures involve training children on the use of context as an aid
in word recognition. The procedures can be introduced at a time
following the introduction of some of the skills associated with
beginning reading. In their 1973 study, Dahl, Samuels and Archwamety
conclude, "Having in a sense replicated studies on two different
populations, retardates and normals, with positive results shown in
both studies, it would seem advisable to recommend these procedures
be used as part of the regular reading curriculum" (p. 72).

This report is the result of expanding the materials by Arch-
wamety and Samuels (1973) and the field testing of these materials
by several EMR classroom teachers. As the initial step in developing a curriculum supplement for classroom use, the experimental materials were expanded and articulated for use by the classroom teacher. This report is the description of the first field test of Hypothesis/Test training procedures involving EMR classroom teachers.

Statement of Purpose

Often a considerable gap exists between what can be done under laboratory conditions as opposed to natural class conditions. The two years of research on the Hypothesis/Test technique in reading was done under conditions not ordinarily found in a classroom. For example, in one study the teaching was done by Research Assistants under the direction of the principle investigators (Archwamety and Samuels, 1973). In a second study (Dahl, Samuels and Archwamety, 1973) the teaching was done by a Ph.D. student skilled in teaching. There is a strong possibility that the amount of resources and teaching help was greater under these conditions than what one ordinarily finds in the classroom.

The purpose of this study was to gain information in the following areas:

1) further development of the Hypothesis/Test procedure training materials,
2) teacher reactions to materials to aid improvement of the materials,
3) replication of the positive results of laboratory studies when the materials were used by the classroom teachers.
Literature Review

Hypothesis/Test procedures were developed in part from research on the process of word recognition, the relation of word recognition to reading comprehension, and the implications of such evidence for the teaching of reading.

The issue of word recognition can be examined from the evidence for serial, parallel and constructive processing in recognizing a printed word. One point of view holds that word recognition occurs by serially processing the letters in the word, one letter at a time. Another view holds that the word is processed as a whole. A third suggests that recognition is a constructive process and only part of the word is used in recognition.

The way in which a word is recognized is influenced by several variables such as the degree of reading skill of the experimental subject and the characteristics of the experiment. Beginning and fluent readers are probably using different strategies of recognition. Whether the words presented for recognition are in isolation or are in context, whether they are high or low frequency words, whether they are easily legible or not, will greatly affect how the words get processed.

As Bradshaw (in press) points out, "A consideration of the serial/parallel distinction in word recognition shows that processing must be serial at some point, whether in handling the line-fragments that make up a letter, the letters that make up a cluster, the syllable or the word. Again, even the most radical parallel model demands serial processing beyond some level whether it is the word,
phrase, or sentence. However, one must consider the task requirements, its difficulty, the experience and the purpose of the reader.\textit{Archwamety and Samuels (1973, p. 1) have said, "Rather than argue that there is but one way to recognize a word, it is probably more correct to argue that the mode of recognition is determined by degree of reading skill and by the demand characteristics of the reading task." This point of view was expressed more than a half-century ago by Huey (1908, p. 81) who said, "The more unfamiliar a sequence of letters may be, the more the perception of it proceeds by letters. With increase of familiarity, fewer and fewer clues suffice to touch off the recognition of the word or phrase, the tendency being toward reading in word-wholes. So reading is now by letters, now by groups of letters or by syllables, now by word-wholes, all in the same sentence sometimes, or even in the same word, as the reader may most quickly attain his purpose."

\textbf{What is the Unit of Word Recognition?}

Sperling (1963) presented subjects with a random matrix of letters followed immediately by a patterned mask. The numbers of letters reported increased linearly with the duration of the matrix; one letter recognized for approximately every 10 msec to a limit imposed by memory. Together with Scharf, Zamansky and Brightbill (1966) where the masked recognition threshold for familiar 5-letter words was roughly 90 msec, evidence was provided for serial processing of letters in a word at the icon level.

Stewart, James and Gough (1969) found that the time between presentation of a word and the beginning of the response increased
steadily as word length increased. McGinnies, Comer and Lacey (1952) selected words which were of approximately equal familiarity but which differed in length and presented them to S's with a tachistoscope. They found that the Visual Duration Threshold was related to word length.

Kamil and Pearson (1972) presented target words which varied in length. In addition these target words had inflected endings such as _s, _ed, and _ing which also varied in length. On the first exposure trial response latency was related to the number of letters in the root of the word. On later trials response latency was not related to word length. This evidence is supportive of serial processing for the first trial, but supportive of wholistic processes for successive trials. The addition of any one of the three inflected endings increased the amount of time required for recognition by an amount approximately equivalent to the addition of one letter on to a stem. This finding regarding inflected endings also supports wholistic processing of well-learned units.

In 1885 Cattell conducted the first of two important studies which were to be used to support reading educators who favored whole word methods of teaching reading. In the first study skilled readers were asked to read a short selection from Gulliver's Travels. They were also asked to name letters equal in number to the number of words in the selection. Cattell found the time required to read an equal number of words and letters was approximately the same and concluded that the subjects were recognizing words as a whole much as though they were responding to a single letter unit. In a
subsequent study using the fall tachistoscope which he had invented, Cattell found that a short word could be responded to as rapidly as a letter.

Although word shape is not necessarily the same as whole word recognition there is some evidence for wholistic processing in Pillsbury's 1897 study in which words containing incorrect letters were presented tachistoscopically to subjects. He found that subjects were able to correctly identify such words and concluded that the subjects were using word shape as the cue to recognition. There was also some evidence that the first few letters of the word were seen more clearly since the misprints were most often detected near the beginning of the word. Criticisms of this conclusion are first that the words were in context and the letters were not aided by a context. Second, it is possible that the subject was using only a word fragment rather than using the whole word for the recognition response.

Work by Broerse and Zwann (1966) indicates that the beginning letters in a word contain the most information as to its identity. Anderson and Dearborn (1952) refer to work by Zeitler in 1900 in which tachistoscopic experiments showed that capital letters and letters extending above and below the line were reported more correctly than others. Zeitler called these "dominant letters."

Goldscheider and Muller (cited in Anderson and Dearborn, 1952) added to Zeitler's list the initial letter of a word as being an important partial cue used in word recognition. Both reports
express the point of view that the function of these dominant letters is not to preserve word shape as a cue, but rather to elicit the pronunciations of the whole word.

Variables Influencing Word Recognition

There are individual differences in strategies used in word recognition. Samuels and Chen (1972) demonstrated that adults recognized tachistoscopically presented words as can children. It was shown that: a) adults demonstrated more faster partial perceptions of a word in the absence of total recognition; b) adults had a better ability to utilize cues such as first and last letters and word length; c) adults showed a greater willingness to alter incorrect hypotheses as to the identity of a word. This evidence in part accounts for the adults' generally faster word recognition.

Smith (1971) theorized that fluent readers use different strategies than beginning readers in recognizing words or in reading a passage. He states that the fluent reader first uses syntax to predict a word and needs only a minimum of visual cues from the printed word to confirm this prediction. The fluent reader uses the printed word to confirm his meaning predictions. In contrast the beginning reader used the surface printed representation to arrive at meaning.

Experimental variations influence word recognition. Tulving and Gold (1968) showed subjects a meaningful context which varied in the amount of information. Following the context, subjects
were then immediately flashed a target word. Speed of recognizing the target word was related to the amount of information contained in the prior context.

According to Tulving, a stimulus word (target) contains a certain amount of information. The subject has to abstract this amount of information before the word can be identified. Some of this information (in some cases all) can be supplied by the relevant preexposure context. Under certain conditions, the preexposure context may contain so much information that the subject needs little or none of the information from the target word to identify the word.

Rouse and Vernis (1963) and Samuels (1969) experimented with a preexposure context of only one word. This word could be an associate of (or relevant to) the target word, or a nonassociate of (or irrelevant to) the target word. Recognition time of the target was faster after the associate word was shown than after the nonassociate word was shown.

Stimulus variables affect the process of word recognition in that characteristics of the stimulus words may affect how fast they can be recognized. The Thorndike and Lorge (1944) word frequency list has been the impetus for research in word frequency. Solomon and Howes (1951) conducted tachistoscopic experiments in which words with high frequencies were recognized faster than words with low frequencies. Again in 1952 Solomon and Postman demonstrated that words with high frequencies were recognized faster than words with low frequencies.
Broadbent (1961) summarized theories attempting to explain this phenomenon of word frequency effect. Broadbent used the term "sophisticated guessing theory" in describing his account for the word frequency effect. Neisser (1967) termed a similar theory "fragment theory." The leading proponent of the "sophisticated guessing theory" or "fragment theory" is Newbigging. According to Newbigging (1961), when a word is presented at a short duration only a few letters of a fragment of the word is seen by the subject. This fragment may be common to a number of words. The subject guesses the word with bias toward the word of greatest frequency of occurrence which incorporates the seen fragment. If the stimulus is a low frequency word, the guess will be wrong and the experimenter has to increase the duration of the stimuli shown. "Word frequency effect" is obtained in this way.

Models in Reading and Word Recognition

The processes involved in word recognition can be viewed in relation to a model of reading. In 1971 Kling reviewed 8544 articles on the reading process in an attempt to come upon the most promising models describing how we read. Williams (1971) grouped existing models into 5 categories:

1. Taxonomic models in which reading behavior is intuitively broken down into several skills.

2. Psychometric models in which reading is analyzed into several independent skills (factors) through the use of the technique of factor analysis.
3. Psychological models which can be further divided into three subcategories: behavioral, cognitive, information processing. Behavioral models describe reading as a process in which an appropriate verbal response is associated with a verbal stimulus through reinforcement. Proponents of this model are Skinner (1957) and Staats (1962). Cognitive models are exemplified by Gibson (1970) who describes learning to read as passing through several phases. As the child becomes more skilled in reading he learns to use the "structural principles" to read in larger, more efficient units. Information processing models are exemplified by Venezky and Calfee (1970) and Smith (1971). In the Venezky and Calfee model two simultaneous processes occur during reading: integration of stimuli already scanned, and forward scanning. The forward scanning locates the largest manageable unit and goes on to the next while the information in the unit is integrated. Smith's model contends that letter identification, word identification, and identification of meaning are all based on feature analysis.

4. Linguistic models which have grown from research in linguistics. One example is Goodman's (1970) model in which the reader decodes from the graphic stimulus not to speech, but directly into deep structure.

5. Transactional models which are exemplified by Rosenblatt (1969). She describes the quality of the experience that the reader is living through, under the stimulus of the text, as the goal of reading. The process is an active two-way relationship between the reader and the text.
Of the five categories, only the psychological models deal explicitly with the problem of how a word is recognized in reading. Williams (1971) states that most of the models are too comprehensive:

Rather, I would like to see us turn our attention to certain limited areas and attempt to refine certain notions that at this point need sharpening. We need "partial" models that are specific, rigorous and testable. Samuels (1971) three-stage model of the recognition of flashed words provides an example. The output of the model is well-specified, the processes are carefully described, and data in support of the model are presented. (p. 158)

Samuels' (1971) three-stage model of the recognition of flashed words referred to above is called the Hypothesis/Test Model of word recognition. The model has been revised (Samuels and Chen, 1972) and now has four stages:

Stage 1: (information use) Information from the reading material already read (or, in Tulving's experiment, the preexposure context) is utilized, e.g., Father the green ______.

Stage 2: (hypothesis making). Information from the reading materials (or preexposure context) as well as knowledge of the structure of English is used to formulate hypotheses, i.e., make predictions of what the next word (or target word) will be, for example, Father cut the green ______ (next word could be emerald, grass, money, plant, etc.).
Stage 3: (test) The hypothesis is tested using new information gathered from partial perception of the next word. Information used to test the hypothesis may be a letter, group of letters, or word shape, e.g., Reader sees letters "gr" which match the word "grass."

Stage 4: (accept/reject) If the next information matches one of the predicted words, the hypothesis is accepted and recognition is rapid. If the new information does not match any of the predicted words, the reader must engage in careful time-consuming visual analysis to recognize the word. The process of word recognition described in the Hypothesis/Test Model is similar to the process of speech perception first advanced by Halle and Stevens (1959) who called this process analysis by synthesis. The essence of the process is that the listener generates guesses as to what a speaker will say and then compares the hypothetical signal with the real ones, i.e., the one produced by the speaker. The perception of speech is achieved even though the listener does not receive the speech signal clearly or in totality. This work is supported by Miller, Heise, and Lichten (1951). They found that words auditorially presented in context and with noisy background were more correctly identified than words auditorially presented in isolation with noisy background.

While the Hypothesis/Test Model has proven useful in accounting for the role of context in word recognition, it has one major problem. The amount of time necessary to generate a prediction is in the neighborhood of 200 ms (Posner and Boles, 1971). Since it only takes about 250 ms or less to recognize a word, this model does not
account for the high speed recognition responses of fluent readers reading meaningful material. Since the Hypothesis/Test procedures are too slow for what goes on in actual reading, the best we can say for it is that it may account for intermediate levels of reading and for tachistoscopic recognition. To account for the actual reading of fluent readers the Hypothesis/Test Model requires further refinement.

**Cloze**

The cloze technique is a procedure in which words are deleted at random or according to some predetermined pattern and the reader or listener is asked to fill in the missing words. The technique was introduced some 20 years ago by Wilson L. Taylor as "a new tool for measuring readability." Investigations using cloze procedure in the past have focused on three areas: cloze as a measurement device of language variables, cloze as a measure of readability, and cloze as a measure of comprehension.

Kennedy and Weener (1973) reported the use of the cloze technique with below average third grade readers in individualized training sessions. The positive effect of cloze visual training found in this study is in agreement with the earlier research done by Best (1971), who also improved the reading comprehension of underachieving readers by training them with the cloze procedure.

A study by Kingston and Weaver (1970) made no attempt to measure an increase in reading proficiency, but rather used cloze procedure with culturally disadvantaged first graders as a predictor of first
grade reading achievement. They found that cloze procedures can be used with first grade subjects and do predict effectively. Kingston and Weaver used a variety of cloze formats including any word cloze, multiple choice, lexical cloze, and aural-reading cloze, random, and every nth deletions. These techniques were used in conjunction with a language experience approach to beginning reading. This study is the most energetic application of cloze procedure in a real instructional situation.

The few studies which have given support to the use of cloze as a means of instruction have in common a real attempt to adapt the cloze procedure to instructional situations. There also seems to be some evidence that deletion of lexical elements is superior to an every nth word system which tends to correlate most with IQ. Perhaps the most striking overall feature of the research on using cloze as an instructional device is the lack of consistent findings. There does seem to be some direction for further research. Future studies should employ lexical element deletions, actively teach strategies for making the closure, use more sophisticated design and measures, and perhaps use subjects relatively less experienced in reading.

**Automaticity**

A concept which must be described in this discussion of word recognition in relation to reading comprehension is that of automaticity. Samuels (1973) has stated that in order to have fluent reading with good comprehension, the student must be brought beyond accuracy to automaticity in decoding.
The need for "automatic habits" in reading is not entirely a new idea. Huey (1908, p. 104) wrote, "Perceiving being an act, it is performed more easily with each repetition... to perceive an entirely new word... requires considerable time and close attention... repetition progressively frees the mind from attention to details, makes facile the total act, shortens the time, and reduces the extent to which consciousness must concern itself with the process." In Fries's (1963) book Linguistics and Reading one can find statements about the importance of automatic habits, but the term "automaticity" is not defined, nor are there explanations of how these automatic habits are developed and measured. While there is a research literature on automaticity in the psycho-motor domain, there is virtually nothing in the verbal learning-reading domain.

To appreciate the power of "automatic decoding" as a psychological process, it is necessary first to discuss the limits of human attention. A quarter of a century of research on attention has led to the conclusion that the brain acts as a single channel processor. This means that at any given moment, attention can only be at one place at a time. If two sources of information are presented simultaneously to a person, each of which demands attention for its processing, the individual finds he cannot process both simultaneously. The individual must choose between them. This dilemma has been described as the "cocktail party problem," a situation one encounters at a party where there are a number of interesting conversations going on at the same time and competing for one's attention. Several choices are available to a person faced with competing sources of information. One choice
involves attending solely to one conversation and ignoring the other sources. The other choice involves attention switching. The individual may be able to follow two or more conversations by rapidly switching attention back and forth. However, at any moment, one's attention can only be on one conversation at a time. The fact that the brain acts as a single channel communication device and can only be attentive to one information source at a time poses important limitations on the beginning reader with regard to comprehending what was decoded. What is remarkable about automatic processes is that they enable a behavior which formerly required attention to be run off without the services of attention. This is tantamount to putting a plane on automatic pilot, thus freeing the pilot to direct his attention to other things which demand it. While numerous behaviors such as tying one's shoe lace or riding a bicycle can be developed to levels which enable them to be performed without attention, it appears that the one important area of human behavior which, regardless of the amount of practice one gets, cannot be performed automatically is that of comprehending language. To comprehend visually or auditorily presented language requires the services of attention.

One can define "automaticity" as follows: A behavior is automatic when it can be performed without attention. Under ordinary circumstances, walking is an automatic behavior. However, when the ground is icy, attention must be used to prevent falling. Another way to approach the problem of defining "automaticity," according to LaBerge (1973) is to consider two tasks which at the unskilled stages
could not be performed simultaneously. Two such behaviors which cannot be performed together at low levels of skill development are sight reading music while at the piano and shadowing speech. After training, if both tasks can be performed simultaneously, at least one of them is automatic. Highly skilled piano players can sight read music and shadow speech. In this case, it is the piano playing which is automatic.

Another example involves automobile driving, since there are interesting parallels between it and reading. At the beginning stages of driving, the student finds that the mechanics of operating the car are so demanding of attention that he finds it difficult, if not impossible, to comprehend conversation while he is driving. Once the student becomes a skilled driver, the mechanical aspects of operating the car get done with no attention, leaving him free to focus attention on processing conversation at the meaning level. Only when some danger signal occurs is the driver forced to direct sustained attention back to the vehicle. When sustained attention is brought back on the mechanics of driving, he finds it impossible to comprehend conversation.

Just as in the examples from the psycho-motor area, where we find important changes taking place as one progresses from unskilled to skilled stages of performance, we find similar changes occurring in reading. At the beginning stages of learning to read, the student’s attention is focused upon the decoding aspects of the task. Since processing information for meaning also requires
attention, as long as the reader's attention is on decoding he cannot comprehend what he has read.

The fluent reader, unlike the beginning reader, is able to decode automatically without the services of attention. Thus, he is able to attend to processing meaning at the same time that he is decoding. Only when a new word appears is the reader's attention directed back to the task of decoding. Once the decoding problem for the new word is solved, the reader's attention can be brought back to processing meaning.

As mentioned earlier, the beginning reader cannot easily comprehend what he has decoded because attention is not available for processing meaning. Unskilled readers can access meaning, however, by rereading a passage several times. The first few readings bring the printed material to the phonological level where it is as if the student were "listening" to it rather than reading it. Once this point is reached, the student is then able to switch attention to deriving meaning from what he has decoded. Teachers who are aware of how difficult it is for beginning readers to access meaning often allow their students enough time to read a passage silently several times before testing their comprehension or asking them to read aloud. This procedure allows the student enough time and trials to switch attention to comprehending the material.

Method

Subjects and Design

The Minneapolis Public Schools gave the writers a list of EMR teachers who would probably be willing to cooperate in such a study
Three classrooms were selected to use the materials with an age range comparable to the Archwamety and Samuels (1973) study. Another 3 were selected with comparable age ranged to serve as comparison classrooms. Three classroom teachers used the materials for approximately one half hour a day for fifteen weeks and three classrooms were used as comparison groups. The average IQ of the children was 72, with a range of 53 to 86. The average age was 11.3 years with a range of 9.2 to 13.2 years. There were 40 children in the 3 classrooms using the materials and 29 children in the classrooms used for comparison. A nonequivalent control group design was used. Because the subjects were not randomly assigned we have pretest data.

The first year of field testing the Hypothesis/Test Model training program followed an eclectic approach to the evaluation of the materials. The roles of this evaluation were to 1) examine the feasibility of this partial reading program, 2) improve the operation of the materials in a classroom setting, and 3) assess the skill achievement of children using these materials. The schedule of data collection and evaluation activities was as follows:

October, 1973 Consent obtained from principles and 6 EMR teachers in the Minneapolis Public Schools to participate in the field testing.

November, 1973 Introduction of teachers to the materials and purposes of the field testing through individual meetings with the teachers.

December, 1973 Pretesting in all 6 classrooms using the tests from the Archwamety and Samuels (1973) study.
January-April, 1974

Program implementation in the classrooms, weekly observations in the three classrooms using the materials and weekly conferences with the teachers using the materials.

May, 1974

Posttesting, using tests from Archwamety and Samuels (1973).

**Description of the data collection instruments or techniques**

The instruments used in the pretests and posttests to help determine student achievement were:
1. A test of word identification in context which was designed to test the child's ability to identify unknown target words in compelling and ambiguous context;
2. A modified cloze test which was one measure of reading comprehension;
3. A test of the seven component skills being taught in the program.

These tests were given individually by the evaluator and a research assistant before and after the units were completed by the classroom teacher. Further description of the tests follows later in this section.

Feedback from the teachers was obtained in scheduled weekly conferences which included questions concerning:
1. What lessons were covered during the week;
2. Completeness of instructions and materials included in the daily lesson plans;
3. Suggestions for improvement in lesson procedures;
4. Conflicts with other reading lessons used by the teacher;
5. Enjoyment of the materials;
6. Confidence in using the materials.

Also in the weekly conference the teachers were asked for comments and concerns about the materials. Occasional observations (about one a week) provided information on whether the materials were used as well as on teacher and student behaviors during the lessons.
Further suggestions were incorporated into the remaining units of the program to be completed or taken note of for incorporation into the program revision. A suggestion was judged to be worthwhile after discussion of the suggestion with the teacher and evaluator. Improvement of the units was the primary focus of the evaluation and continual teacher feedback was the process used for suggestions for program improvement. Student achievement was considered, but only as one aspect of the evaluation. Student achievement results are of use for information about whether the units for the classroom teacher are as effective as the original experimental set-up with research assistants doing the teaching. This evaluation could be described as a pilot test of the program materials and final judgments as the "worth" of this program are not included in the evaluation. The aims of this evaluation are improvement of the materials and feasibility of the program.

**Description of instructional materials used**

Hypothesis/Test: The following subskills and the methods of instruction were derived from a task analysis of the Hypothesis/Test model of word recognition (Archametey and Samuels, 1973).

Materials used by the teachers during the 15 weeks on instruction were based on the following subskills:

- Training in the ability to construct a word given an initial sound. Students were given lessons of the following nature:
  - **Stimulus situation**: Teacher says, "Tell me a word starting with the sound /p/.
  - **Response situation**: Child gives a word starting with the sound /p/.

-
2. Training in the ability to tell the starting letter of a word just heard. Students were given lessons of the following nature:

   **Stimulus situation:** Teacher asks, "What is the first letter in the word 'girl'?"

   **Response situation:** Student gives the name of the initial letter in 'girl'.

3. Training in the ability to visually recognize the initial letter of a word heard. Students were given lessons of the following nature:

   **Stimulus situation:** Teacher says, "What is the first letter in the word 'boy'." Children choose from the printed letters b c d r.

   **Response situation:** Student responds by indicating the letter b.

4. Training in the ability to use auditory context to predict words that could logically follow in a sentence without hearing the initial sound hint as to what the word to follow the context might be. Students were given lessons of the following nature:

   **Stimulus situation:** Teacher says, "My mother cooks in the _______."

   **Response situation:** Students predict the missing word.

5. Training in the ability to use auditory context to predict word(s) that could logically follow in a sentence hearing just the initial sound hint as to what the word(s) to follow the context might be. Students were given lessons of the following nature:

   **Stimulus situation:** Teacher says, "The cat ran after the _______."

   **Response situation:** Students predict what the missing word might be.
6. Training in the ability to use visual context to predict word(s) that would logically follow in a sentence without seeing the initial letter hint as to what the word(s) to follow the context might be. Students were given lessons of the following situation:

**Stimulus situation:** Teacher shows the following in printed form: The children open the ______.

**Response situation:** Students are asked to read and predict the word in the blank. Teacher tells students word that cannot be read.

7. Training in the ability to use visual context to predict word(s) that could logically follow in a sentence seeing the initial hint as to what the word(s) to follow the context might be:

**Stimulus situation:** Teacher shows the following in printed form: The girl ate the b______.

**Response situation:** Students are asked to read this and predict the word in the blank.

8. In addition to the subskills, vocabulary used in cloze stories was included in the materials. Daily vocabulary practice was a part of each lesson along with subskill practice.

The concept of automaticity was explained to the 3 teachers using the materials, and they were asked to encourage speed, not just accuracy, in all responses of the lessons based on the skills described above.

Tests of Student Achievement

**Modified Cloze Test**

The material used was an 8" by 11" piece of paper on which a passage was printed. Twenty words in this passage were deleted except for their initial letters. The passage was as follows:
Bobby and Johnny went to the z____. They saw k____, ch____, t____, l____, birds, m____ and many other animals. Bobby liked b____ animals b____ Johnny liked s____ animals. Bobby bought some f____ to feed the sea lion. Johnny gave c____ to squirrels to eat. When it was getting d____, the boys put their bicycles in the g____. Then, they drank some w____. Their sister, Patsy, was sewing some clothes. She had with her some th____ and a n____. After supper, everyone was in the living room singing. Mother was playing the p____, and father was playing the v____. It was a j____ day.

The following instructions were used:

Today we are going to p____ some guessing games. You and I are going to read this (point to the experimental material): When we come to a blank such as this (point to the first blank) you have to guess what the word in the blank should be. The word you guess must start with the letter you see in the blank. Any questions? Let's start.

The child read the passage and the tester helped the child read any word the child could not read. When the child came to a blank, if the blank was at the end of the sentence the tester asked, "Can you guess what the word in the blank is?" No help was given on the target word. If the blank was not at the end of the sentence, the child was asked to read on to the end of the sentence. Then the tester asked, "Can you guess what the word in the blank is?"

1. started with the letter appearing in the blank,
2. was grammatically correct,
3. made sense,
he received one point for that blank. If he failed in any of these three requirements he received zero points for that blank.
Word Identification in Context

The materials used were two sets of ten 5 x 8 index cards. On each card was typewritten a sentence. The last word of the sentences was underlined in red and called the target word. The words before the target word were called the context. During the test the tester sat opposite the child at a table. He or she covered the context of a sentence with a blank index card. Only the target word was exposed. The tester then asked the child, "Can you read this word?" If the child read the target word correctly, he was ready for the next trial. If the child could not read it, the tester said, "Let's try it this way." The tester then exposed the context and said, "Read from here," pointing to the first word of the context. The tester helped the child read any word the child could not read until the child came to the target word. If the child could then read the target word by himself, he received one point. If he could not, the tester asked, "Can you make a guess?" If the child guessed correctly, he received one point. If he guessed incorrectly or refused to guess he received zero points. Then, the child was ready for the next trial or sentence. The tester gave no help on the target word. The child was given one point only if he gave the same word as appeared in the sentence.

Sentences in the first set were intuitively designed such that the context seemed so compelling that the target word could hardly be any other word. These ten sentences were:

1. The Apollo astronauts went to the moon.
2. It is dark at night.
3. When it is dark, we turn on the light.
4. Mother sewed her clothes with a needle.
5. Father pounded the nail with a hammer.
6. I saw the smoke coming-out of the chimney.
7. That sick man was sent to the hospital.
8. I put the ring on my finger.
9. Mr. Smith has two sons and three daughters.
10. The loud noise the sky makes when it rains is called thunder.

Sentences in the second set were designed such that the context might suggest some other word than the target word. The context in the second set was not as compelling as that of the first. The ten sentences in this second set were:

1. That fierce animal is a tiger.
2. There are fish in the lake.
3. The car is making a funny sound.
4. At the hospital, there are nurses.
5. The teacher told us a joke.
6. At Christmas, Alex received a lot of gifts.
7. Father keeps his tools in the garage.
8. We get the news from the newspaper.
9. Mother likes to drink coffee.
10. We drink milk.

Hypothesis/Test Component Skills Test

Part I Auditory Test

1.1 Testing the use of context to predict word(s) that could logically follow in a sentence.

Instruction:
We are going to play a game. I'll say something and you'll guess what I am going to say next. "On Sundays, we go to the ______." What word do you think comes next?
"I put on my ice ______." "Johnny drinks a cup of ______."

(real test)
1. Father called to ______.
2. We write on the ______.
3. Mother cleaned the ______.
4. The dog bit the ______.
5. Cats like ______.
6. Astronauts can walk on the _____.
7. After supper we wash the _____.
8. When it rains I put on a _____.
9. The mailman delivers _____.
10. At noon we eat _____.

1.2 Testing the use of context and partial perception to predict word(s) that could logically follow in a sentence.

Instruction:
We are going to play another game: I'll say something and you'll guess what word comes next. I'll help you guess. Tommy sat on the "ch____." What word do you think comes next?

"Father took me to the picture show. We saw a good m____."  
"He is my fr____."

(real test)
1. Father hit the b____.
2. In the classroom, we listen to the t____.
3. Babies don't walk, they cr_____.
4. Can you stand on your h____?
5. I found a shiny new p______.
6. Father pounds nails with a h______.
7. A garden has fl______.
8. On my birthday, mother may bake a c____.
9. During the summer, I can ride by b______.
10. In bed, I cover up with a bl______.

1.3 Testing the construction of a word starting with the given sound.

Instruction:
We are going to play another game. I'll say a sound. You'll tell me a word starting with that sound. "p______." Can you tell me a word starting with that sound?
"g____"
"sh____"

(real test)
1. b____  
2. d____  
3. t____  
4. th____  
5. ch____  
6. b____  
7. br____  
8. gr____
1.4 Testing the ability to catch the initial sound of a word.

**Instruction:**
We are going to play another game. I'll tell you a word. You'll tell me what letter of the alphabet it starts with. "boy" (flash card). What letter does it start with? "girl" "man"
Now you'll have to do it without a flash card. "dog"
What letter does it start with? "cat" "bird"

(Real test)
1. grass
2. bread
3. blue
4. thunder
5. chair
6. bat
7. dog
8. table
9. hat
10. water

Part IIWritten Test

2.1 Testing the use of context to predict word(s) that could logically follow in a sentence.

**Instruction:**
Read this sentence and tell me what word comes next.
I like to eat ____
She has a ____
Mother called to the ____

(Real test)
1. I open the ____
2. The pen is on the ____
3. Do you know my ____?
4. I walk to the ____
5. I can paint the ____
6. Firemen ride in ____
7. Batman is on ____
8. A jet lands at the ____
9. I can sign a ____
10. The warmest season is ____
2.2 Testing the use of context and partial perception to predict word(s) that could logically follow in a sentence.

Instruction:
Read the sentence and tell me what word comes next.
We hit the b_____.
The dog chased the c_____.
There are a chair and a t_____.

(real test)
1. He likes to eat p_____.
2. Mary likes to s_____.
3. She bought a c_____.
4. He saw a m_____.
5. The house was built of br_____.
6. Some airplanes are j_____.
7. I have a new pair of sh_____.
8. Some children are boys and some children are g_____.
9. The mouse ate the ch_____.
10. At the circus, I saw a cl_____.

Part III Cross-modal test (auditory + visual)
Testing the ability to recognize the initial letter given a word audiotorially.

Instruction:
Listen to the word "boy." What letter does it start with? (show m b x i)
Listen to the word "girl." What letter does it start with? (show c f g z)
Listen to the word "cat." What letter does it start with? (show c m g b)

(real test)

Listen to the word: What letter does it start with?
1. grass g p q y
2. bread b d p q
3. blue d p b q
4. thunder t f h l
5. chair c e o u
6. bat q p b d
7. dog q b d p
8. table f h t l
9. hat n h u y
10. water v w m n
Results

In all tests the analysis was a pretest to posttest comparison using the student's test. The posttest scores of the classrooms using the materials were compared to their pretest scores. The posttest scores of the classrooms not using the materials were compared to the pretest scores for those classrooms. Gain scores were used in the analysis rather than analysis of covariance because gain score analysis is easier to interpret and the assumptions are less stringent. The number of subjects on the different measures varies slightly due to absenteeism on testing days.

Pretest and posttest comparisons of the seven component skills of the classrooms using the units follow on Tables 1 and 2. Pretest and posttest scores for the same time period from the classrooms not using the units are shown on Tables 3 and 4. Table 1 compares pre and post reaction times for the classrooms using the units. Table 2 compares pre and post number correct for each component skill for the classrooms using the units. Table 3 compares reaction times for the classrooms not using the units before and after the time period in which the units were used. Table 4 shows the number correct for the classrooms not using the units before and after the time period in which the units were used.

It can be seen in Table 1 that the reaction time gains on the component skills made by the classroom using the units were significant at the .05 level on five of the seven skills, and in Table 2 one notes the number correct gains were significant at the .05 level on three of the seven skills. The significant gains made by the classroom not using
<table>
<thead>
<tr>
<th>Component Skill</th>
<th>Pretest</th>
<th>Posttest</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>SD</td>
<td>N</td>
</tr>
<tr>
<td>1.1 Auditory Prediction Without Aid</td>
<td>1.35</td>
<td>.53</td>
<td>40</td>
</tr>
<tr>
<td>1.2 Auditory Prediction With Aid</td>
<td>1.06</td>
<td>.47</td>
<td>40</td>
</tr>
<tr>
<td>1.3 Constructing a Word Given a Sound</td>
<td>1.96</td>
<td>.78</td>
<td>40</td>
</tr>
<tr>
<td>1.4 Telling First Letter Given Word</td>
<td>1.49</td>
<td>.74</td>
<td>40</td>
</tr>
<tr>
<td>2.1 Visual Prediction Without Aid</td>
<td>1.18</td>
<td>.57</td>
<td>40</td>
</tr>
<tr>
<td>2.2 Visual Prediction With Aid</td>
<td>1.16</td>
<td>.71</td>
<td>40</td>
</tr>
<tr>
<td>3. Visual Recognition First Letter Given Word</td>
<td>1.75</td>
<td>.61</td>
<td>40</td>
</tr>
</tbody>
</table>

Note—time is given in seconds.
Component Skills Abbreviations

1.1 Given a spoken sentence with one word deleted completely, S can insert (orally) an appropriate word.

1.2 Given a spoken sentence with one word deleted except for the word's initial letter sound, S can insert (orally) an appropriate word.

1.3 Given a sound S can construct (say) a word starting with that sound.

1.4 Given a spoken word, S can state the initial letter.

2.1 Given a visually presented sentence with one word deleted completely, S can insert (orally) an appropriate word.

2.2 Given a visually presented sentence with one word deleted except for the word's initial letter. S can insert (orally) an appropriate word.
### TABLE 2

NUMBER CORRECT RESPONSES FOR CLASSROOMS USING MATERIALS

<table>
<thead>
<tr>
<th>Component Skill</th>
<th>Pretest</th>
<th>Posttest</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>SD</td>
<td>N</td>
<td>X</td>
</tr>
<tr>
<td>1.1 Auditory Prediction Without Aid</td>
<td>9.73</td>
<td>0.55</td>
<td>40</td>
<td>9.90</td>
</tr>
<tr>
<td>1.2 Auditory Prediction With Aid</td>
<td>9.15</td>
<td>0.74</td>
<td>40</td>
<td>9.40</td>
</tr>
<tr>
<td>1.3 Constructing a Word Given a Sound</td>
<td>6.58</td>
<td>1.22</td>
<td>40</td>
<td>6.74</td>
</tr>
<tr>
<td>2.1 Visual Prediction Without Aid</td>
<td>8.90</td>
<td>1.15</td>
<td>40</td>
<td>9.28</td>
</tr>
<tr>
<td>2.2 Visual Prediction With Aid</td>
<td>8.50</td>
<td>1.75</td>
<td>40</td>
<td>9.08</td>
</tr>
<tr>
<td>3. Visual Recognition of First Letter</td>
<td>8.60</td>
<td>1.80</td>
<td>40</td>
<td>9.40</td>
</tr>
</tbody>
</table>
TABLE 3

MEAN REACTION TIME FOR CLASSROOMS NOT USING MATERIALS

<table>
<thead>
<tr>
<th>Component Skill</th>
<th>Pretest</th>
<th></th>
<th>Posttest</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>SD</td>
<td>N</td>
<td>X</td>
<td>SD</td>
</tr>
<tr>
<td>1.1 Auditory Prediction Without Aid</td>
<td>1.26</td>
<td>0.39</td>
<td>29</td>
<td>1.33</td>
<td>0.62</td>
<td>29</td>
</tr>
<tr>
<td>1.2 Auditory Prediction With Aid</td>
<td>0.99</td>
<td>0.29</td>
<td>29</td>
<td>0.92</td>
<td>0.28</td>
<td>29</td>
</tr>
<tr>
<td>1.3 Constructing a Word Given A Sound</td>
<td>1.87</td>
<td>1.04</td>
<td>29</td>
<td>1.93</td>
<td>1.03</td>
<td>29</td>
</tr>
<tr>
<td>1.4 Telling First Letter Given Word</td>
<td>1.39</td>
<td>0.62</td>
<td>29</td>
<td>1.40</td>
<td>0.61</td>
<td>29</td>
</tr>
<tr>
<td>2.1 Visual Prediction Without Aid</td>
<td>1.52</td>
<td>1.05</td>
<td>29</td>
<td>1.29</td>
<td>0.71</td>
<td>29</td>
</tr>
<tr>
<td>2.2 Visual Prediction With Aid</td>
<td>1.31</td>
<td>0.55</td>
<td>29</td>
<td>1.18</td>
<td>0.58</td>
<td>29</td>
</tr>
<tr>
<td>3. Visual Recognition of First Letter Given a Word</td>
<td>1.91</td>
<td>0.93</td>
<td>29</td>
<td>1.85</td>
<td>0.71</td>
<td>29</td>
</tr>
</tbody>
</table>

Note—time is given in seconds.
### TABLE 4

CORRECT RESPONSES FOR CLASSROOMS NOT USING MATERIALS

<table>
<thead>
<tr>
<th>Task Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{X}$</td>
<td>SD</td>
<td>N</td>
<td>$\bar{X}$</td>
</tr>
<tr>
<td>1.1 Auditory Perception Without Aid</td>
<td>9.52</td>
<td>0.69</td>
<td>29</td>
<td>9.69</td>
</tr>
<tr>
<td>1.1 Auditory Perception With Aid</td>
<td>8.86</td>
<td>1.03</td>
<td>29</td>
<td>9.24</td>
</tr>
<tr>
<td>1.2 Constructing a Word Given a Sound</td>
<td>5.66</td>
<td>1.57</td>
<td>29</td>
<td>6.45</td>
</tr>
<tr>
<td>1.3 Missing First Letter of a Word</td>
<td>8.78</td>
<td>1.66</td>
<td>29</td>
<td>8.52</td>
</tr>
<tr>
<td>1.4 Visual Perception of First Letter</td>
<td>8.07</td>
<td>1.92</td>
<td>29</td>
<td>8.43</td>
</tr>
<tr>
<td>1.5 Visual Recognition of First Letter</td>
<td>8.14</td>
<td>1.75</td>
<td>29</td>
<td>8.45</td>
</tr>
</tbody>
</table>
the units was on only one of the seven skills for reaction time and there were no significant gains when the number correct was measured.

The summary results of the modified cloze test are shown on Table 5. The analysis of the modified cloze test shows that the classrooms using the units filled in significantly more acceptable words than the classrooms not using the units when each group was compared pre to post.

Table 6 shows the pre to post comparisons on recognition of words in context for the classrooms using the materials and for those not using the materials. As can be seen from Table 6, both groups of classrooms made significant gains on this measure.

Both sets of classes gained significantly on the total number of target words correctly identified. Classrooms using the materials made significant gains on both ambiguous and compelling target words. The classrooms not using the units gained significantly on the words in ambiguous context. Both sets of classes identified significantly more target words in the compelling context.

The weekly observations and conferences with the teachers using the units provided the following information:

1) The materials were easily implemented into the existing reading program. The materials in this program did not conflict with the usual reading activities.

2) The teachers used the materials approximately 1/2 hour per day for 15 weeks.

3) Teachers individualized the materials in an appropriate manner according to teacher style and skill needs of the children. For example,
Table 5

NUMBER OF ACCEPTABLE WORDS IN BLANKS FOR THE MODIFIED CLOZE TEST

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Posttest</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLASSROOMS USING MATERIALS:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>SD</td>
<td>N</td>
<td>X</td>
</tr>
<tr>
<td>Pretest</td>
<td>11.69</td>
<td>3.62</td>
<td>36</td>
<td>13.64</td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CLASSROOMS NOT USING MATERIALS:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>SD</td>
<td>N</td>
<td>X</td>
</tr>
<tr>
<td>Pretest</td>
<td>10.42</td>
<td>3.75</td>
<td>26</td>
<td>11.27</td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 6

NUMBER CORRECT FOR WORDS PRESENTED IN COMPELLING OR AMBIGUOUS CONTEXT

Classrooms using materials:

<table>
<thead>
<tr>
<th></th>
<th>( \bar{X} ) pretest</th>
<th>( \bar{X} ) posttest</th>
<th>SD difference</th>
<th>N</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compelling context:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>posttest-pretest</td>
<td>8.35</td>
<td>9.19</td>
<td>1.38</td>
<td>37</td>
<td>3.68</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>Ambiguous context:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>posttest-pretest</td>
<td>7.19</td>
<td>8.57</td>
<td>2.27</td>
<td>37</td>
<td>3.70</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>Total: posttest-pretest</td>
<td>15.54</td>
<td>17.16</td>
<td>3.16</td>
<td>37</td>
<td>4.26</td>
<td>&lt;.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>( \bar{X} ) compelling posttest</th>
<th>( \bar{X} ) ambiguous posttest</th>
<th>SD difference</th>
<th>N</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posttest-compelling-ambiguous</td>
<td>9.19</td>
<td>9.57</td>
<td>1.59</td>
<td>37</td>
<td>2.58</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Classrooms not using materials:

<table>
<thead>
<tr>
<th></th>
<th>( \bar{X} ) pretest</th>
<th>( \bar{X} ) posttest</th>
<th>SD difference</th>
<th>N</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compelling context:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>posttest-pretest</td>
<td>7.82</td>
<td>8.32</td>
<td>1.71</td>
<td>28</td>
<td>1.55</td>
<td>&lt;.10</td>
</tr>
<tr>
<td>Ambiguous context:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>posttest-pretest</td>
<td>5.93</td>
<td>7.46</td>
<td>1.64</td>
<td>28</td>
<td>4.95</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>Total: posttest-pretest</td>
<td>13.75</td>
<td>15.79</td>
<td>2.50</td>
<td>28</td>
<td>4.32</td>
<td>&lt;.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>( \bar{X} ) compelling posttest</th>
<th>( \bar{X} ) ambiguous posttest</th>
<th>SD difference</th>
<th>N</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posttest: compelling-ambiguous</td>
<td>8.32</td>
<td>7.46</td>
<td>1.85</td>
<td>28</td>
<td>2.44</td>
<td>&lt;.05</td>
</tr>
</tbody>
</table>
one of the classrooms made the vocabulary practice into a game called "Bust-it" and the teacher requested that the short cloze stories be written on tagboard rather than individual dittoed sheets because they would be easier for her class to use.

4) The teachers enjoyed the materials and asked to keep them.

5) The children seemed to enjoy the units. Whenever observed, the children in all three classes were cooperative and used the materials as instructed.

The teacher at School 1 had spent much of the school year on letter sound--letter symbol correspondence prior to the use of the Hypothesis/Test materials. Her class was able to begin immediately with vocabulary practice along with subskills 2 through 7. She used the "Bust-it" vocabulary game in which the class was divided into two groups for practice. The children were eager to play whenever the class was observed; the teacher said she continued with this approach to vocabulary throughout the 15 weeks "because the kids really enjoy it." She requested that the cloze stories be written on tagboard for her class since it was easier for them to attend when they did not have individual dittoed cloze stories. The children attended to the lessons whenever observed.

The teacher at School 2 followed closely the directions given in the materials. After completion of the 15 weeks she stated she "will always do cloze because the children like it and it holds their attention." She practiced letter sound--letter symbol correspondences with the children in her class who were not skilled in this ability. During this time the other children in her class were doing repeated readings.
on stories below their reading level. (Repeated readings are briefly explained in the discussion section.) Vocabulary practice was done in small groups with the more mature readers helping the younger readers in reviewing words learned. When observed one day, the teacher asked the children to come sit for vocabulary and one child turned to the observer and said "The [vocabulary] cards are really fun." The children needing practice with letter sounds and letter symbols played letter bingo and letter domino for extra practice. These games are part of the instructional materials.

School 3 spent the longest time on sound-symbol correspondences since the children's abilities in this area needed improvement. Vocabulary practice was done in groups of two, with one child who knew a particular set of words well helping another child needing more practice. The teacher would remind the children occasionally about the importance of automaticity: "You have to be able to say the words real fast." The cloze stories were presented on tagboard as with School 1 because it was easier for the children to follow.

Discussion

The purpose of this study was to field test materials based on the experimental materials described by Archwamety and Samuels (1973) and Dahl, Samuels and Archwamety (1973). The specific purposes were as follows: 1) to further develop the Hypothesis/Test procedure training materials, 2) to observe teacher reactions to materials to aid improvement of the materials, and 3) to discover whether the positive results from laboratory studies were replicated when classroom teachers used
the materials. Information was collected through weekly conferences and observations of the classes using the materials and through pre and posttests of the competent skills. The observations and conferences gave information concerning the difficulty of implementation of the materials and suggestions for classroom management of this program. The pre and post tests were used for information about student achievement.

As noted in the results, it can be said that the classrooms using the experimental materials had acquired a better set of competent skills comprising the Hypothesis/Test word recognition process than the classrooms not using the materials.

Gains were measured on the basis of pre to post score differences since the classes were assigned intact to either those using the materials or those serving as a comparison group. As can be seen in Table 7, the ages and IQ's of the classes were not significantly different. Analysis of covariance was not used since the pre to post comparisons for each group of classrooms yielded sufficient information for this initial field test of materials. A more rigorous experimental design would perhaps be more appropriate after this consideration of the feasibility of the materials.

The results of the component skills test do not say unequivocally that the materials and instructional program produced higher gains on the subtests by the classrooms using the materials. Perhaps the three teachers using the materials were superior to those who served as comparison teachers. However, the component skills gains were more often
### Table 7

**Age and IQ Comparisons for Both Groups of Classes at the Time of Pretesting**

<table>
<thead>
<tr>
<th></th>
<th>Classrooms using materials</th>
<th>Classrooms not using materials</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IQ</strong></td>
<td>X 73.31 SD 7.08 N 32</td>
<td>X 71.79 SD 8.48 N 24</td>
<td>.24</td>
<td>.809</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>X 11.17 SD .95 N 40</td>
<td>X 11.46 SD 1.09 N 27</td>
<td>-1.11</td>
<td>.273</td>
</tr>
</tbody>
</table>

*Note: age is in years to the nearest/hundredth.*
significant for the classrooms using the materials and demonstrated that the skills in the Hypothesis/Test Model of word recognition can be taught. Further field testing is needed to help clarify the possibility of an interaction between teacher skill and style and the materials and instructional program.

Another point which should be noted is that the scores of many children in both sets of classes were near mastery even in the pretests. Referring to Tables 2 and 4 it can be seen that a possible 10 correct on all subtests, except for 1.3 which had a possible 8 correct, did not leave much room for improvement on these skills. The children in the classes had learned many of these component skills through the usual reading instruction. It is the suggestion of the writers that the instruction with the Hypothesis/Test materials may be more appropriate with a younger age group of EMR children. With a wider range of pretest abilities on the subtests the question of what level of entering skills is most benefitted by these materials could be answered.

The results shown in Table 5 indicated that the classrooms using the experimental materials showed a significant gain on the modified cloze test from pre to post testing. The classrooms not using the materials did not show a significant gain in correct responses during the same time period.

Table 6 shows that both groups of classrooms made significant gains on this measure of identifying unknown words in either a compelling or an ambiguous context. The skills measured by this task were learned by both sets of classes.

The question which must be asked in further field testing is...
whether or not the Hypothesis/Test procedures would aid the comprehension and word recognition strategies of a more inexperienced reader. The study shows that the Hypothesis/Test model of word recognition can be taught as a set of component skills and that these component skills can be taught to EMR children by their classroom teacher.

The study showed that an instructional program to help EMR children to predict unknown words as an aid to comprehension has been constructed. The approach to teaching reading in this program is to encourage children to make good predictions on target words, using information from context and partial visual cues from the target words. The procedures are straightforward, simple, do not require teacher retraining, are inexpensive and easily evaluated.

In the LaBerge and Samuels (1973) theoretical paper on automaticity, there is a section on awareness when one is accurate in a response as opposed to when one responds automatically. The authors conjectures that the subject probably has greater knowledge of component skills and features of the stimulus when he is accurate than when he is automatic. At automatic levels the subject is often unaware of how he is processing stimuli. This discussion is relevant to the training of Hypothesis/Test recognition strategies. The subject is trained in the component skills used in Hypothesis/Test recognition, but we wish him to become automatic. Consequently, it might be desirable for him to lose his awareness of what these component processes are since he should be very fast in his responses.
There is still a more important reason for wanting the subject to be automatic with Hypothesis/Test procedures. If the procedure requires the services of attention, then it is probable that there will be interference with the comprehension process. In both the LaBerge and Samuels (1973) and the Samuels (1973) papers on automaticity, there is the suggestion that in fluent reading the decoding must be done automatically, i.e., without attention, for attention to be on deriving meaning from the decoded material. Consequently, it is most important that Hypothesis/Test procedures be run off automatically so that the subject can understand what he is reading.

In this first year of field testing the feasibility of the Hypothesis/Test instructional materials has been explored. In all three schools teacher reactions to the materials were favorable. The problems encountered concerning classroom management during the use of the materials centered on the fact that all classrooms using the materials contained a wide range of reading abilities. Because of this, much of the instruction had to be individualized. For example, the children needing letter sound-letter symbol instruction could not proceed to the cloze stories until this subskill was mastered. Also, children who had difficulty learning the vocabulary used in the cloze stories were unable to read the cloze stories independently. This management difficulty was handled smoothly in School 3 by having the more mature readers repeatedly read stories below their reading level. Each story was read four times. These "repeated readings" seem to be in accordance with the theory of automaticity (Samuels, current study). The children
willingly participated in this task and this freed the teacher to work with those children needing other subskill instruction.

Another problem related to the wide range of reading abilities within each classroom was the 15 week instruction time. In one classroom using the materials more time spent on letter sound--letter symbol correspondences and vocabulary practice would have been appropriate in order for the children to become more "automatic" in their responses. However, the instructional time limit in this study required that all units be completed in the 15 weeks.

Perhaps the most outstanding feature of the materials is that the cloze procedure in the Hypothesis/Test materials is novel to most classrooms and therefore seems to be of high interest. Both auditory cloze and visual cloze can be done individually or in groups and can be applied to any reading materials used in the classroom.

Synthesizing the information gained from the teacher conferences and classroom observations, it would seem that two major issues should be considered in the revision and further field testing of these materials. One is that it is necessary to assess the impact of Hypothesis/Test materials with readers not so advanced in the subskills of these materials. This would involve using younger EMR children. The second issue is that more cloze stories should be incorporated into the later units and perhaps the focus of the materials should be on children who have just mastered the subskills of letter sound--letter symbol correspondences. When much of the 15 weeks is spent on sound--symbol correspondences (which is a subskill in most reading programs) the impact of the Hypothesis/Tests would seem to be lessened. It would seem that
Hypothesis/Test training would be appropriate for those children who have mastered sound-symbol correspondences and are ready for a strategy to aid in comprehension.

Further field testing of these materials should investigate for whom the training is most valuable, that is, the age at which the instructional program provides the most benefit. In this study the general age group studied by Archamety and Samuels (1973) was used, but their skills upon entering this program were superior to those reported by Archamety. The next evaluation of these materials should include a younger EMR age group. Also to be included in the next study is the tachistoscopic word recognition tests, in which the speed of recognition of a known word is measured.

The materials are currently being revised according to the information provided by this first year's feasibility study of these materials. The program will be field tested another year with more emphasis on the teacher effects in relation to the materials, with a younger age group included, and with more tests of word recognition and comprehension.


Archwamey, T. and Samuels, S. A mastery based experimental program for teaching mentally retarded children word recognition and reading comprehension skills through use of hypothesis/test procedures.


Bradshaw, J. L. Three interrelated problems in reading: a review (in press).


Rate of information processing in visual recall and methodological considerations.


*Theory of Reading.* Berkeley:

Read: *Reading Research Quarterly,* 1969, 111.


Finkston, A. and Weaver, W. Feasibility of cloze techniques for teaching and evaluating culturally disadvantaged beginning readers. Research and Development Center in Early Educational Stimulation, University of Georgia, 1970.


Taylor, A. The use of immediate recall in rapid comprehension. Memory, 1951, 1, 155-175.


Technological Reports

University of Minnesota Research, Development and Demonstration Center in Education of Handicapped Children

(Place of publication shown in parentheses where applicable)


And education of the learning-disabled child in the Soviet Union. Research Report #78.


Initial field test and feasibility study of the hypothesis/word recognition pro-


The universality of self-generated verbal mediators as a means of enhancing memory pro-


L. Donnelly. Development trends in the generation and utilization of associative rela-

ures. The acquisition of specific reference in the linguistic system of a deaf child of


early maternal linguistic environment of normal and Down's Syndrome (Mongoloid) language


Other reading skills acquired by low readers taught in regular classrooms using clinical research Report #35. May 1972.


Prevalence of school learning & behavior disorders in a longitudinal study population.


w. & s. jansen. Syntactic elaboration in the learning & reversal of paired-associates by


The effects of varied levels of verbal mediation on the learning and reversal of paired
Deficiency, 1971, 76, 60-67. Study II: American Journal of Mental Deficiency, 1971, 76,

s. & t. jones. Effectiveness of model guidance, modeling, trial and error learning for
girls with institutionalized behavior. Research Report #4. June 1970. (" ill-
1973, 19, 59-61.)

Effects of social and social distractors by moderately retarded institutionalized children.


Diagnostic Pedagogy, International Congress on Deafness. Stockholm, August 1970; also
American Instructors of the Deaf Annual Convention, St. Augustine, Florida, April 1970.

christ. Systematic scoring of ranked distractors for the assessment of Piagetian reasoning

-sequence hypothesis: Individual differences in the development of school related spatial