DOCUMENT RESUME

ED 098 516 CS 001 427

AUTHOR Guthrie, John T.
TITLE Identification and Instruc-ion of Children with Reading Disability. Second Annual Report to the Spencer Foundation.
INSTITUTION Johns Hopkins Univ., Baltimore, Md. John F. Kennedy Inst.
SPONS AGENCY Spencer Foundation, Chicago, Ill.
PUB DATE Jul 74
NOTE 253p.; Some pages of document have marginal reproducibility

EDRS PRICE MF-$0.75 HC Not Available from EDRS. PLUS POSTAGE
DESCRIPTORS Reading; *Reading Ability; Reading Achievement; *Reading Difficulty; *Reading Processes; *Reading Research; *Reading Skills; Reading Tests; Retarded Readers

ABSTRACT

This current annual report describes the progress made in the following aspects of the organizational plan; the identification of a model of reading which includes a description of subskills that may be present in skilled readers and are acquired by most normal readers, the construction of criterion referenced tests to measure each of these subskills, a comparison of normal and disabled readers in terms of their performance on the component subskills, and development of curricular procedures for disabled readers. The contents of the report include; "Overview of the Project," which discusses the organizational plan; "Model of Reading," which directs the reader to the model presented in the first annual report; "Measurement of Reading Subskills," which discusses the measuring instruments used, the areas of reading evaluated and the results obtained from the use of the measuring instruments; "Comparison of Normal and Disabled Readers," which discusses the learning hierarchies for good and poor readers, and an analysis of the comparative skills of normal and disabled readers; "Summary of Cognitive Deficiencies of Disabled Readers," which reviews several studies related to the cognitive deficiencies of disabled readers; "Affiliated Projects," which presents several research projects conducted in collaboration with the staff of the Spencer Foundation Project; and "Dissemination of Findings." (WR)

SECOND ANNUAL REPORT

to the

SPENCER FOUNDATION

On A Grant Entitled

"IDENTIFICATION AND INSTRUCTION OF CHILDREN WITH READING DISABILITY"

Awarded To:

John T. Guthrie
Director of Special Education
John F. Kennedy Institute
Assistant Professor of Pediatrics
The Johns Hopkins University

July 1, 1974

Table of Contents

ANNUAL REPORT
To The
SPENCER FOUNDATION

John T. Guthrie

Overview of the Project

At the outset of this report, a review of the general strategy for the conduct of the present project will be presented. As indicated in the first annual report, there are four components to the organizational plan for this investigation. First, a model of reading must be identified which includes a description of subskills that may be present in skilled readers and which are acquired by most normal readers. Second, criterion referenced tests must be constructed to measure each of these component subskills. Third, normal and disabled readers should be compared in terms of their performance on the component subskills. Analyses of cognitive deficiencies of disabled readers should also be conducted with other procedures. Research paradigms drawn from the literature of visual information processing and psycholinguistics may be used. This information will amplify the uniquenesses of disabled readers which are illustrated by their performance on the tests of reading subskills. Fourth, from the newly discovered information regarding the nature of the acquisition of subskills and the cognitive deficiencies of disabled readers curricular procedures for disabled readers should be developed. These methods may be pilot tested and rigorously evaluated in field tests. The current annual report will describe the progress made in each of these aspects of the organizational plan. Also included is a description of the collaborative research efforts in which the staff from the project are involved and a description of the dissemination of our findings and results.

Model of Reading.

Since the model of reading which is used for the framework of this investigation was presented in the first annual report, a description of the model will not be presented here. Little change has been made in our perception of the model and we have continued to use it as a foundation for the construction of tests and analysis of reading subskills.

Measurement of Reading Subskills.

Progress in the measurement of reading subskills has been substantial this year. In the first year of the project a preliminary version of the Kennedy Institute Phonics Test (KIPT) was developed. The measure was administered to a group of disabled readers and two contrast groups of normal readers, as described in the previous report. Based on that investigation substantial revisions and improvements on the KIPT have been made. Of the original fifteen subtests on the KIPT, eight have been retained. One of the original subtests was eliminated due to low reliability. Four were eliminated due to redundancy with other subtests as indicated by their high intercorrelations and similar degrees of difficulty; and two were eliminated due to their low value for instructional purposes. The revised version of the KIPT has ten subtests. The new measures are designed to assess capability with consonant vowel combinations both in the production and recognition modes.

The final list of subtests is: 1) Whole Word Production-Special Rules, 2) Nonsense Word Production, 3) Long Vowel Word Production, 4) Short Vowel Word Production, 5) Consonant Vowel Production, 6) Letter Sound Production, 7) Letter Naming, 8) Nonsense Word Recognition, 9) Consonant Vowel Recognition, and 10) Initial Letter Sound Recognition. Each subtest represents a carefully specified domain of reading stimuli. These domains, which were vague in the first version, have been clarified. The domains are described in detail in the preliminary manual of the KIPT which is included as appendix 1 of this report.

The domains are defined by phonological categories. Inspiration for these categories was drawn from the computer charting of phonological rules by Venezky (1970). For example, in the Short Vowel Word Production Subtest all the words contain the vowel sounds that occur in the following words: bat, bet, bit, but, dot. There is no restriction on the letters that may be used to represent the vowel sounds in these words

That is "head" may occur since it contains the same vowel sound as "bet". In the Long Vowel Word Production Subtest vowel sounds of the words include the sounds which occur as follows: tape, meet, kite, tune, boat. There is no restriction of the letters that may be used to represent these sounds. For example, "might" could occur since the vowel sound is the same as that of "kite". Thus regularities of grapheme-phoneme correspondence are not considered in this taxonomy of reading stimuli.

The current version of the Kennedy Inst" 't Phonics Test contains four parallel forms. Each of these forms has ten subtests with items for the forms being drawn randomly from the domains that are specified in the preliminary manual. In addition to the generation of these forms, the preliminary manual was written this year. This manual includes the rationale for the test, description of how the subetsts were constructed, the limited available technical information and detailed directions for administration of the measure.

To my gratification I believe that the KIPT is worth publication. Support for that belief comes from the 25 or 30 requests that we have received from large school districts for systematic use of the test. I have initiated a contact with Consulting Psychologist Press for negotiations of the publication of the KIPT. The psychometric support for this test, however, is not sufficient yet and publication now would be premature. I have requested $5000 from the Consulting Psychologist Press to support a research assistant one-half time for a year, who will collec. .e necessary data to establish the reliability of the revised subtests and the equality of difficulty in the four parallel forms. The establi.hment of validity of a criterion-referenced measure is a controversial topic. Although this criterion-referenced test has self-evident face validity, efforts will be made to establish empirical documentation for the relevance of this test to reading diagnosis and instruction. A copy of the available materials including the preliminary manual and the four parallel forms are enclosed as appendices 1 and 2.

It is intriguing that the subtests of the KIPT are rank ordered in difficulty in a very consistent manner for normal and disabled readers alike. At present a doctoral candidate at Johns Hopkins University is conducting a dissertation on the hierarchical analysis of reading subskills using this measure. He is conducting a Guttman scale analysis and an experimental transfer test of the hierarchical nature of these subtests. Since that work is currently in progress, it will be reported in next year's annual report.

Since the decoding skills in the KIPT appear to form such a clear hierarchy, it was decided that the possible presence of a comprehension hierarchy should also be explored. To examine whether there is a learning hierarchy in reading comprehension for normal and disabled readers a lengthy investigation was undertaken. An attempt was made to measure component skills of comprehension and to determine whether these skills are arranged in a hierarchy of difficulty. The tasks used to explore this issue were: 1) Sight Vocabulary. This test included forty items with four alternatives. The subject's task was to identify the alternative which has the closest meaning to the stem. 2) Sentence Reading Comprehension. The measure used to assess this skill was the maze task which was previously used in the Spencer Projects (Guthrie, 1973b). 3) Oral Reading. This task consisted of requiring subjects to read isolated words aloud. 4) Sentence Listening Comprehension. This task was parallel to the Sentence Reading Comprehension except that the materials were read aloud to the children and their answers were given in a response mode similar to the Sentence Reading Comprehension. 5) Sentence Reading Comprehension (Questions). This task attempted to assess the children's ability to answer questions (of four categories) which were posed for sentences that were presented in paragraph format. These five tasks were given to a normal group of 24 second graders and a group of 24 disabled readers who attend the Kennedy School. The groups were matched on I.Q. and reading level but were obviously different in chronological age.

From a measurement viewpoint the efforts were successful. The reliability with the KR-21 formula for all of the tasks for normal readers ranged from .85 to .98. One exception was the Sentence Reading Comprehension (Questions) task which had a reliability of .73, probably due to its reduced number of items. The reliability for disabled readers on these tasks ranged from .83 to .97. The exception in this task was also the Sentence Reading Comprehension (Questions) task which had a reliability of .79. All of the tasks were moderately correlated with Gates-MacGinitie Comprehension Test performance. These correlations indicate that while the subtests appear to be related to reading they are not correlated at such a high level that one may assume that they are measuring the same ability as the standardized reading test. Intercorrelation among tasks for normal readers ranged from .66 to .86; and the intercorrelation for disabled readers ranged from .21 to .90. Although most of these intercorrelations were substantial, they were not so high as to indicate that the tasks were simply measurements of the same underlying skills.

Measurement of reading subskills in reading comprehension requires more than the establishment of reliability and low-to-moderate subtest intercorrelation. If the concept of a hierarchy of subskills is to be successful, differential difficulty of the tasks must be established and the criterion of positive transfer between successive subskills must be met. Essentially this effort requires the comparison of normal and disabled readers in the array of tasks which have been described previously. Consequently, the efforts to measure components of the reading process and the comparison of normal and disabled readers on these components were both included in this investigation.

Comparison of Normal and Disabled Readers.

The learning hierarchies for good and poor readers were compared with two methods. First, the difficulty of each of the tasks for both groups was assessed. For normal readers the rank order was such that Oral Reading was the easiest, Sentence Listening Comprehension and Sentence Reading Comprehension were significantly more difficult and did not differ from each other; and Sentence Reading Comprehension (Questions) and Sight Vocabulary did not differ from each other and were significantly more difficult than the previous set of skills. For poor readers the hierarchy was substantially different. Sentence Listening Comprehension, Oral Reading and Sentence Reading Comprehension were similar to one another. Sight Vocabulary was significantly more difficult and Sentence Reading Comprehension (Questions) was significantly more difficult than all the other skills.

A Guttman Scaling Analysis of the learning hierarchies for the two groups was undertaken. This analysis revealed that normal readers had a Guttman scale consisting of: Oral Reading, Sentence Listening Comprehension, Sentence Reading Comprehension and Sentence Reading Comprehension (Questions). The reproducibility of this scale was .94, which is substantial. For poor readers the hierarchy consisted of: Sentence Reading Comprehension, Oral Reading, Sight Vocabulary and Sentence Reading Comprehension (Questions). The reproducibility of this scale was 1.00. This high reproducibility means that the order of difficulty was consistent for all subjects on all tasks. It is noteworthy that the learning hierarchies of the two groups are different both in terms of the skills which are included and in terms of the order of skills which both groups have in common. Among poor readers Sentence Reading Comprehension (Questions) was a uniquely difficult task for about one-third of the group indicating that answering questions poses a markedly high challenge for disabled readers. This hierarchy information will be used as a basis for construction of curriculum sequences to be used in the Kennedy School and to be pilot tested in the coming year. Further details of this investigation are presented in the paper entitled, "Learning Hierarchies in Reading Comprehension" which is included as appendix 3 in this report.

Comparisons between normal and disabled readers on a variety of tasks which are pertinent to the reading model have been conducted. In one study, (Guthrie & Seifert, 1973), normal and disabled readers were given measures of oral reading proficiency and sentence reading comprehension. In summary, the results of this experiment were that poor readers were worse than normal readers when they were matched on Sight Vocabulary. The poor readers were inferior on both Oral Reading and comprehension, although their inferiority in comprehension was more marked than their difficulty in Oral Reading. Normal children who read aloud with only moderate proficiency understood at least some of the material. However, poor readers who made even a few errors while reading the words in the passage aloud did not comprehend the material at the most elementary level. It is clear that assessing, teaching and monitoring reading comprehension is particularly important for poor readers. Details of this experiment are reported in a paper entitled, "The Maze Technique for Assessing and Monitoring Reading Comprehension", which has been accepted for publication in The Reading Teacher, a publication of the International Reading Association.

Also included in the manuscript described in the previous paragraph is a description of how the maze technique may be used by classroom teachers to assess the progress of children in reading comprehension. Suggestions are given about how to redirect comprehension teaching in a manner that will be sensitive to the children's learning abilities. The publication of this article represents a substantial effort to disseminate the results of our work to classroom teachers and other practitioners in the field of reading who may be able to make use of our results. This paper is included as appendix 4.

As an analysis of the comparative skills of normal and disabled readers in visual identification of word constituents (letters), phonemic extraction of sounds from words and semantic coding of words was conducted by a post-doctoral fellow, Rick Steinheiser, who worked under my direction. One result of his investigation was that semantic targets (words) were identified faster than letter targets which were in turn identified faster than phonological targets while reading normal text materials. These results confirm the findings of other investigators, (Cohen, 1970; Smith & Haviland, 1972). Data now point to the conclusion that, in reading, the word is encoded as a "Gestalt" from which specific graphic information may then be extracted, along with the segmentation into pronounceable sequences. Rather than constructing a word from its letters it would appear that the word is the unit of analysis from which more refined information (semantic, graphic, phonological) is segmented. Our results tend to support the contention of Smith (1973) who suggests that "meaning can be extracted from sequences of written words independently of their sounds, and in fact meaning must be comprehended if sounds are to be appropriately produced". (Page 78).

An intriguing relationship between performance on the search tasks in disabled readers and a reading matched control group was found. The disabled readers were more proficient in identification of single letters than this contrast group, but they were significantly less proficient in the extraction of sounds. In addition, the disabled group and this normal control group were similar in the identification of semantic targets. These facts combine to indicate that disabled readers rely less heavily than normal readers on the sounds of words for the extraction of meanings. Disabled readers rely more heavily on graphemic structure of words for derivation of meaning than normal children. In other words disabled readers adopt a compensatory strategy in reading. Since they cannot decode words into sounds as readily as normal children, they learn how to derive meaning directly from the graphological cues that are available. However, this compensatory strategy does not permit disabled readers to read as well as children their same age, since their performance on the semantic search task in this study was significantly inferior to that of an age matched control group. Information on this study is contained in a manuscript entitled, "Scanning Times Through Prose and Word Strings for Various Targets by Normal and Disabled Readers", by Steinheiser & Guthrie. This paper has been accepted for publication in Perceptual and Motor Skills and is enclosed as appendix 5 in this report.

An investigation of the visual perception and phonological coding skills of normal and disabled readers was conducted by Rick Steinheiser with my guidance in the past year. In this study very simple tasks were used with the sensitive measure of reaction time as the dependent variable. Children were presented two words with the vowel sounds underlined such as head and bead. The words were presented simultaneously. Children were asked, in one condition, to say whether the underlined portions looked the same. This is a simple task of matching visual forms. In another condition the children were asked to indicate whether the underlined portions of the words sounded the same. They pressed a button to indicate whether they thought the items were the same or different in both conditions. The visual processing (physical match) of disabled readers was relatively good, being better than that of reading match controls and inferior to that of age match controls. However, the phonological coding (sound match) of disabled readers was markedly deficient. Their reaction times were more than twice as long as those of reading match controls which were in turn more than twice as long as those of age match controls. The conclusion of this study was that while poor readers appear to receive the information visually with the relatively good efficiency their ability to extract sounds from the accurately perceived letters is remarkably poor. This study is included in a manuscript entitled,"Adequate Visual Encoding Followed by Poor Phonological Encoding As A Source of Reading Disability", and it is enclosed as appendix 6 in this report.

Summary of Cognitive Deficiencies of Disabled Readers.

Our studies are accumulating to give us an intriguing picture of the cognitive deficiencies of disabled readers. The reading model that we are relying on (following Mackworth, 1971) contains the four stages: 1) visual representation, 2) coding, 3) storage in memory, and 4) response production. Previously we outlined the "system model" which contends that these four components are acquired in a dynamic, interdependent manner rather than an independent fashion (Guthrie,1973a). Development of one of the processes facilitates development of the others.

Not only does our previous paper (Guthrie, 1973a) provide evidence for this view, but the recent studies also support it. The investigations reported in immediate preceding paragraphs all found that disabled readers were inferior to their age matched controls on all tasks. This would be predicted if one assumes that a deficiency in one basic process reduces development in other processes and leads to a generalized retardation in basic cognitive operations that are relevant for reading. However, comparison of disabled readers with children matched with them on reading level provides an indication of relative strengths and weaknesses of cognitive processes. With these comparisons one or more outstanding cognitive deficiencies can be observed. What we observe by comparing disabled children to reading matched controls is that they are adequate on the first component of the reading process i.e. visual representation. This evidence derives from two separate investigations by Steinheiser and Guthrie. Other data in the field (Noelker & Schumsky, 1973 and Senf & Feshbacn, 1970) do not contradict this position.

The second phase of the reading process, i.e. coding, may be divided into phonological and semantic coding operations. Disabled readers appear to be deficient in phonological coding. In both of the Steinheiser and Guthrie studies and in the Guthrie and Seifert (1974) study the ability of disabled readers to derive sounds from printed materials was demonstrated to be deficient. Phonological coding is a complex operation and further study is needed to determine whether this deficiency is derived frcm deficiencies of phonemic segmentation, memory for auditory materials or grapheme-phoneme correspondence skills.

Although our data is not complete on this issue, semantic coding of disabled readers appears to be adequate. In the Steinheiser and Guthrie Study (1974) the disabled reader's ability to identify semantic targets in passages, was equal to that of reading matched contrast groups. In the Guthrie and Seifert (1974) study, the sight vocabulary of disabled readers was actually superior to that of reading matched controls. Needless to say, we do not know the units that are involved in semantic coding. Such coding may occur at the word level, two word level, phrase level, etc. It is clear that the semantic interpretation of sentences is deficient for disabled readers (Guthrie, 1973b). Analysis of this deficiency will require complex investigations. We do not know the roles of the many variables of semantic and syntactic cues in sentences, retrieval of knowledge and information from memory that interacts with the information in the sentence, and the effect of phonological distortions in reading on meaningful processing of the sentences.

The fourth component of the process which consists of reading response production and expectancy responses have not been fully analyzed. Preliminary indications are that reading response production is normal since subjects could identify periods in one of the first tasks in Steinheiser and Guthrie paper (See appendix 5), and since the speech of disabled readers is usually tested to be normal.

Development and Evaluation of Instructional Principles.

A fundamental purpose of this project is to develop instructional concepts and principles for the remediation of children with reading disability. The fourth stage of the organizational plan for the study consists of the development and evaluation of instructional principles. Based on a model of reading, tests which are developed to measure reading subskills in the model, and information that is derived from the comparison of normal and disabled readers, our instruction should be more effective than traditional approaches.

The primary area of instructional development has been the construction of a hierarchical mastery procedure for teaching decoding skills. The need for carefully engineered instruction in this domain issues from our evidence that phonological decoding of disabled readers is severely deficient. The teaching strategy that appears to be effective is based on the hierarchy of decoding subskills that are defined by the subtests of the Kennedy Institute Phonics Test. Our teaching strategem is simple. We administer the Kennedy Institute Phonics Test to children individually. A profile provides an immediate basis for the initiation of instruction. Skills that the child has mastered above the 90% proficiency level are not taught. The least difficult skill which has not been learned to the 90% proficiency level is the starting point for instruction. For example, a child may have a proficiency profile consisting of Single Letter Sounds, 95%; Consonant Vowel Sounds, 90%; Short Vowel Word Production, 70%; Long Vowel Sounds, 30%; Nonsense Words, 20%; Whole Word Production, 15%. In this case, the teaching would be initiated with short vowel words. This domain of decoding would be taught to proficiency of 90% accuracy. At that point, instruction in long vowel words would be started and would continue until mastery level was obtained.

To provide a preliminary evaluation of the effectiveness of the hierarchical mastery teaching procedure, a comparison between progress of children in the Kennedy School and progress of children in remedial reading programs in public schools was undertaken. In September, 1973 twenty-two children in the Kennedy School were assigned to the hierarchical mastery procedure in decoding. The children were taught in the following manner. The Kennedy Institute Phonics Test was administered to each child at the outset of the school year by the research assistant. The profiles were given to the teachers and children with similar profiles were grouped together. Teaching was initiated on the reading subskill which was the least complex in the hierarchy that had not been mastered (90% proficiency). Children were taught this subskill to mastery level and the next highest skill in the hierarchy was then taught. Each of the subskills was taught in sequence until the hierarchy of ten subskills had been completed. Children were taught for a period of forty minutes per day in teacher-pupil ratios of 1:4 and 1:2. All of the teachers were staff of the Kennedy School who are certified as Reading Specialists and have been trained in this school. The materials used for instruction were constructed by the teachers and include many games to facilitate the acquisition of specific reading subskills. Selected subtests of the KIPT are given periodically, i.e. about every two weeks, in order to assess progress.

Children in the Baltimore County Public School System were compared with the Kennedy School children. There are two schools which specialize exclusively in reading disability that are sponsored by Baltimore County. The twenty-two children in the Kennedy School were matched with twenty-two children in these two public school programs. The mean ages of the Kennedy School children and public school children were 10.21 and 10.27, respectively with standard deviations of 1.232 and 1.231. The reading levels on the Gates-MacGinitie and the Metropolitan Achievement Tests were 2.41 and 2.35 for the Kennedy School and public school respectively with standard deviations of .39 and .32. Thus the age and reading levels of the children did not differ significantly. In a larger investigation we have determined that the I.Q. socio-economic level and sex distribution of the children in these three schools does not differ. The program used in the public school consisted of a teacher-pupil ratio of 1:3 and 1:7 during reading class. The reading period was sixty minutes in length. In addition children received tutoring on a based teacher-pupil ratio of 1:2 for one-half hour approximately three times a week. The methods of instruction were eclectic. Teachers were given information regarding the child's I.Q., reading level, auditory, visual and kinestetic strengths and weaknesses. Teachers used a linguistic approach emphasizing Palo Alto materials, Merrill readers and Lippincott basals. Phonics instruction, word knowledge training, comprehension instruction and modified Orton-Gillingham and Fernald techniques were used depending upon the teacher's ongoing diagnosis.

The results of the comparison were that the reading level in Gates-MacGinitie Vocabulary Test for the Kennedy School children at the end of the year in May, 1974 was 3.29 grade level with a standard deviation of .73. The reading level for the public school children at the end of the year was 2.69 grade level and the standard deviation was .56. The difference between those grade levels is statistically significant ($t = 3.04$, $df = 42$, $p < .01$). The Kennedy School children made a mean gain of .88 years whereas the public school children made a mean gain of .34 years in reading vocabulary. Thus our program produced more than twice as much growth as a model program in an affluent public school with a very similar student population. A more detailed description of this comparison will be forthcoming in the near future. The evidence from this camparison is provocative. The hierarchical mastery approach for remediation of disabled readers in decoding appears to be potent and warrants careful experimentation in the future. We would like to design an educational experiment to validate the principle of hierarchical mastery instruction for reading disabled children. Such research encounters a host of methodological difficulties and will require considerable investment in time and money. However, we are in a position to validate an educational principle and to thereby identify a mode of operation which may be used by mary different professionals in many different educational environments.

A pilot program for instruction in reading comprehension was initiated in September, 1973 with 16 children in the Kennedy School. This program consisted of simultaneously monitoring three skills that are related to reading comprehension. The hierarchical mastery procedure which is used in decoding was also used in this comprehension teaching strategy. The three skills included were: Oral Reading, Sentence Reading Comprehension, and Paragraph Comprehension. It was speculated that these skills would form a hierarchy such that a child would have to be able to read a given passage orally before he could understand the sentences, and he would have to understand each of the sentences before he could answer questions over paragraphs of material. The program was initiated by obtaining descriptive information on each child's reading level at the outset. Performance on three tasks including oral reading, sentence comprehension, and paragraph comprehension were observed for textual materials which were judged to be at an intermediate level of difficulty for each child. For example, a passage at the 2^1 level might be selected. A child would be given an Oral Reading, Sentence Comprehension and Paragraph Comprehension task over subsections of the passage. The objective was to teach him to read orally with 90% proficiency then teach him to comprehend the sentence with 90% proficiency and then instruct him in paragraph comprehension skills to the level of 90% proficiency.

The difficulty of the material and the skill to be taught were determined by identifying the task which the child performed at less than 90% proficiency but for which he had appropriate prerequisite skills. For example, if a child performed 90% proficiency in Oral Reading on a 2^1 passage, 70% proficiency on the Sentence Reading Comprehension and 30% on the Paragraph Reading Comprehension for those materials, he would be instructed in sentence comprehension at the 2^1 level of difficulty. Instruction continues until the 90% proficiency level was reached at which time paragraph comprehension instruction would be intiated. After paragraph comprehension was achieved to 90% proficiency, the child would be given oral reading, sentence comprehension and paragraph comprehension tests at the 2^2 level of difficulty. If he was proficient in each of these skills at this level, tests at the next level would be given. However, if one of his skills was lower than mastery level, the skills were taught and the cycle were repeated. Children in this program had an entering Gates-MacGinitie vocabulary grade equivalent of 3.30 and a Gates-MacGinitie comprehension grade level of 2.75. The group gained a mean of .85 years on vocabulary and .80 years in comprehension. These gains in comprehension are gratifying since previous attempts to influence comprehension as such had not been successful.

There are two reasons for not attempting to evaluate the hierarchical mastery approach to reading comprehension yet. The first is that, in the pilot study, the skills used were based on speculation rather than data. The investigation entitled, "Learning Hierarchies in Reading Comprehension", that was undertaken this year provides a hierarchy for use with this type of teaching strategy. The skill hierarchy for disabled readers observed in the study was: Sentence Reading Comprehension, Oral Reading, Sight Vocabulary and Paragraph Reading Comprehension. In the coming year these four skills will be used as the foundation for instruction. Continuous monitoring of all four skills and hierarchical mastery instruction within this system will be conducted. Since the hierarchy to be used in the coming year is new, a rigorous documentation of its instructional benefits will not be initiated. A second reason for not attempting to document the validity of this instructional strategy is that the methods for measuring paragraph reading comprehension have not been fully validated and verified. Additional basic research in that issue is necessary before instructional implications and programs can be developed.

Affiliated Projects.

A number of research projects are being conducted at the Kennedy Institute in collaboration with the staff of the Spencer Foundation Project. Two of these collaborations have already borne fruit in terms of significant findings and publications. Other collaborations in the more primitive stages also will be reported. First, the genetic study of reading disability conducted by Joan Finucci with the collaboration of Professor Barton Childs, Professor of Pediatrics in Johns Hopkins University, has been completed. This project was conducted partially in the Kennedy School and enjoyed the assistance of Spencer Foundation staff members. In this study eighteen white families each having at least one child with specific reading disability were studied to determine the extent of reading disability within the families. Scores from a battery of reading, spelling, and intelligence tests were used to make classifications of family members as normal or disabled readers. In addition to a standardized reading test, three reading tests were designed specifically for this study to detect adult readers who may have compensated for a reading disability which they more clearly manifested as children. In addition to family members, 72 adults in a comparison group were administered all or part of the test battery to determine criteria for classification as a normal reader.

The male to female ratio among index cases was 15:3. Among the siblings of the index cases 83% of the brothers were affected and 20% of the sisters were affected. Of the 17 fathers examined 10 were classified as affected and among 17 mothers examined 7 were classified as affected. The greater number of affected males than affected females cannot be explained by sex linkage.

A procedure was devised for testing a hypothesis of sex control. This hypothesis poses that reading disability has a dominant manifestation in males and recessive manifestation in females. In addition to accomodating the disproportion in sex ratio of affected individuals and the families in which both parents were normal readers, the sex controlled hypothesis is supported by the fact that better agreement was found between the expected and observed value of affected males for a small (.01) rather than a larger (.20) frequency of a gene in the population controlling reading disability. Further investigation of the genetic component is currently underway with particular emphasis on the issue of testing multi-genetic hypotheses and attempting to explore possible subtypes of reading disability in order to differentiate modes of transmission for these subtypes.

Our neurological analysis of reading disability is also alive. Last year Drs. Preston, Childs and I observed that disabled readers had neurological deficiencies in the left parietal lobe as measured by the Visual Evoked Response Technique. This year parents of children in the Kennedy School have been tested with the Visual Evoked Response Technique. Parents who were classified in Finucci's study as normal or disabled were both tested and a control group of normal parents with normal children was also tested. The intriguing result was that a light flash stimulus did not differentiate among the three groups. However, a stimulus condition which included the presentation of words and the requirement that the subject process words differentiated the groups. The normal control group manifested the largest amplitude in the visual evoked responses from the left parietal lobe. The parents of disabled children who were normal readers had significantly smaller visual evoked responses than the control group. Parents who were disabled readers had significantly smaller responses than parents of disabled readers who were normal in reading ability. The fact that parents of children with reading disability should have a neurological deficiency even though they are normal readers is valuable information. It suggests that this technique may detect individuals who are recessive for the gene or individuals who manifest the phenotype in a neurological but not in a behavioral fashion.

The post-doctoral fellow working in the Division of Special Education, Dr. Anthony Frankfurter, is collaborating with Dr. Dennis Fisher, a psychologist in Aberdeen Proving Grounds, on a study of high speed visual processing in normal and disabled readers. Their hypothesis is that iconic storage of sequences of visual stimuli may be deficient in poor readers. Their investigation is in progress and a full report will be provided at another time.

The Kennedy Institute is fortunate to have Dr. Paula Tallal as a post-doctoral fellow. She received her degree in experimental psychology from University of Cambridge, England and is spending two years here studying dyslexic and aphasic children. She is currently conducting an analysis of auditory processing deficiencies in both aphasic and dyslexics. With her methodology she is able to explore the time parameters of auditory input in sensitive manner. This information may shed light on the origin of the phonological coding deficiencies of disabled readers which were discussed previously in this report. Since her work is also under progress, it will be reported later when it has been completed.

A doctoral dissertation which was conducted in the Kennedy School with my support related to the issue of parental attitudes and learning in disabled readers. The title of the dissertation was "Socio-Economic Status, Parent-Child Rearing Attitudes and Marital Adjustment: The Relationship Between These Factors and the Children's Rate of Response to Special Education for Dyslexia". In this study children's reading gains will be correlated with the parental responses to a "Marital Adjustment Test" and "The Maryland Parent Attitude Survey". This exploration of affective factors in reading disability is a valuable contribution to our information in this field since many teachers believe that marital discord can be both the cause and the result of reading failure in schools. Further investigation of these issues after the completion of this dissertation will be warranted. A report of results will be enclosed at a later period.

Dissemination of Findings.

Dissemination of the findings of our projects has taken place through several avenues. The recent articles by Guthrie in the Journal of Educational Psychology entitled, "Models of Reading and Reading Disability" and "Reading Comprehension and Syntactic Responses in Normal and Disabled Readers" have elicited about four hundred reprint requests particularly from Europe, Canada and Australia. Other articles which are in press include: a manuscript entitled, "Visual Evoked Responses in Normal and Disabled Readers" by Preston, Guthrie and Childs which is to be published in Psychophysiology. The paper entitled, "The Maze Technique for Assessing and Monitoring Reading Comprehension" by Guthrie, Seifert, Burnham & Caplan has been accepted for publication in The Reading Teacher. The basic investigation of cognitive processes by Steinheiser and Guthrie entitled, "Scanning Times Through Prose and Word Strings for Various Targets for Normal and Disabled Readers" which will be published in Perceptual and Motor Skills in 1974.

Two of my associates have read papers at national conventions recently. Mary Seifert, a research assistant on the Spencer Project, wrote and delivered a paper entitled, "Reading

Comprehension Deficiencies in Disabled Readers" for the Annual Convention of the Association for Children with Learning Disabilities in Miami, in 1973. Tim Sheehan, a graduate student working with me, conducted and reported a study at the American Educational Research Association in Chicago, 1974. This paper was entitled, "Perception and Memory in Performance of Normal and Disabled Readers on An Auditory Discrimination Task. (Appendix 7,8). Finally, I am providing consultation on research in reading disability to three prominant organizations including: National Science Foundation, National Institute of Education and The National Institute of Law Enforcement and Criminal Justice, a branch of the Department of Justice. All of these organizations are initiating basic and applied research on reading disability and we are providing assistance in the development of guidelines and review of the proposals.

REFERENCES

Cohen, G. Search times for combinations of visual, phonemic and semantic targets in reading prose. _Perception & Psychophysics_, 1970, 8, 370-372.

Guthrie, J.T. Models of Reading and Reading Disability. _Journal of Educational Psychology_, 1973a, 65 (1), 9-19.

Guthrie, J.T. Reading Comprehension and Syntactic Responses in Good and Poor Readers. _Journal of Educational Psychology_, 1973b, 65 (3), 294-299.

Guthrie, J.T. and Seifert, M. The Maze Technique for Assessing and Monitoring Reading Comprehension. _The Reading Teacher_, 1974, in press.

Mackworth, J.F. Some models of the reading process: learners and skills readers. In F.B. Davis (Ed.) _The Literature of Research in Reading, with Emphasis on Models_. New Brunswick, N.J. School of Education, 1971.

Noelker, R.W. and Schumsky, D.A. Memory for Sequence Form and Position as Related to the Identification of Reading Retardates. _Journal of Educational Psychology_, 1973, 64 (1), 22-31.

Seifert, M. and Guthrie, J.T. Reading Comprehension Deficiencies in Disabled Readers. _Proceedings of International Congress of Learning Disabilities_, 1973.

Senf, G. and Feshback, S. Development of Bisensory memory in culturally deprived, dyslexic and normal readers. _Journal of Educational Psychology_, 1970, 61 (6), 461-470.

Smith, E. and Haviland, S.E. Why words are perceived more accurately than non-words: inference vs. unitization. Journal of Experimental Psychology, 1972, 92, 59-64.

Smith, F. Psycholinguistics and Reading. New York, Holt, Rinehart and Winston, 1973.

Steinheiser, R. and Guthrie, J.T. Scanning times through prose and word strings for various targets by normal and disabled readers. Perceptual and Motor Skills, 1974, in press.

Venezky, R.L. The Structure of English Orthography. The Hague Mouton, 1970.

Kennedy Institute Phonics Test;

Experimental Version

John T. Guthrie & Mary Seifert

Kennedy Institute, Johns Hopkins University

Introduction

The Kennedy Institute Phonics Test (KIPT) was constructed for three purposes: 1) to describe the subskills involved in single word reading, 2) to facilitate the instruction of children who are poor readers, and 3) to provide an instrument that is useful for research on the acquisition of reading, the analysis of reading disabilities and the teaching of reading. The fulfillment of each of the goals requires considerable time, effort and data. Since the measure has not been evaluated from either a psychometric or instructional viewpoint, its success in fulfilling the purposes cannot be determined at this time. However, a brief discussion of the characteristics of the KIPT as a criterion-referenced test and preliminary information regarding the technical qualities of the test is provided. This experimental version of the KIPT is intended to facilitate the development of the test and to assist professionals who are responsible for service programs in reading.

I. Rationale for the Kennedy Institute Phonics Test

A distinction between norm-referenced tests and criterion-referenced tests was drawn by Glaser in 1963. He stated that "criterion-referenced measures depend on an absolute standard of quality, while . . . norm-referenced measures depend on a relative standard." (p.519) Information obtained from a criterion-referenced measure of reading may indicate that a child can read 3 letter short vowel words with 70 percent accuracy; whereas the results of a norm-referenced test may indicate that the child reads at the 2.0 grade level. In the first case, the child's score may be interpreted without reference to the performance of other children, but in the second case the only information available is the comparison of the child to his peers, in terms of grade level.

Considerable controversy exists regarding the nature of criterion-referenced measurement (Kriewall, 1972; Nitko, 1971; Roudabush and Green, 1971; Lyman, 1972; Hambleton and Novick, 1972; Popham, 1971). Since a review of the issues is not possible here, suffice it to say that a point of agreement of all of the authors is that a "criterion-referenced test is any test for which a criterion score is specified in a way that does not depend on the scores of the examinees to whom it will be applied" (Livingston, 1973, pp. 13-14). Using this definition, the KIPT qualifies as a criterion-referenced test since a criterion passing score of 85 percent correct is recommended for each subtest.

The issue of whether the measure may be regarded as a mastery test should be raised. Although at least one author (Wentling, 1973) uses the term mastery test to refer to a criterion-referenced measure of achievement, it is self evident that the two terms are not always synonomous. Perhaps a criterion-referenced measure with a high criterion of proficiency may be termed a mastery test if the universe of items represented by the test is unambiguously defined. In any event, it is tentatively suggested that the KIPT is a mastery test in the sense that performance of a child on a subtest at the level of 85 percent accuracy implies a high level of proficiency on all of the items in the universe represented by the subtest. If a child performs with 85 percent accuracy on the short vowel subtest, it is probable that he has "mastered" all of the short vowel words as defined by the guidelines used to construct the subtest.

II. Construction of Subtests

The KIPT includes four alternate forms, C, D, E, F, each of which includes all of the subtests. Each of the subtests was constructed to represent a sample of items from a finite domain of possible items. This section presents tentative definitions of these domains. The order of presentation of the definitions is intended to simplify the descriptions of the domains and is not similar to the order of administration of the subtests.

Subtest 7 - Letter Naming

A random sample of 20 letters from the 26 lower case letters in the alphabet was used.

Subtest 6 - Letter Sound Production

A random sample of 20 letters was drawn from the population of all letters in the alphabet excluding q and x.

Subtest 5 - Consonant Vowel Production

The items were formed by combining all consonants and all vowels in 2 letter combinations, excluding q. Combinations from this pool were selecte randomly with the exception that items were excluded in which the consonant changes the sound of the vowel as in "ew" or the vowel modifies the most commonly used sound of the consonant as in "ce".

Subtest 4 - Short Vowel Word Production

All words in this domain contain the vowel sounds that occur in the following words: bat, bet, bit, but, dot. There is no restriction on the letters that may be used to represent the vowel sounds in these words. That is "head" may occur since it contains the same vowel sound as "bet". The words in this domain contain all the following single consonant sounds: bet, cat, dig, fig, gab, hat, jab, kin, lit, mit, not, pot, rat, sat, top, tax, vat, wig, yet, zip. In addition, all 2 and 3 letter consonant blends are contained in this domain including: brim, bled, crib, clip, drip, frog, flop, grim, glad, prim, plan, stop, slim, skit, strip, scan, sprig, trip, twist, blend, most, limp, ring, melt, snap, [illegible], bank, scalp, chasm, bent, swim, milk, self, film, rasp, bask, pact. In the subtest, about 50 percent of the items contain consonant blends and 50 percent contain single consonant sounds.

Subtest 3 - Long Vowel Word Production

The vowel sounds of the words in this domain include the vowel sounds of the following words: tape, meet, kite, tune (also cube), boat. There is no restriction on the letters used to repersent these sounds, e.g. "might" could occur since the vowel sound is the same as that of "kite". The consonant sounds include all of the sounds contained in the domain of Subtest 4 with the addition of: consonant digraphs such as: th, ch, sh, gh, ph, thr, wh; the soft c and g as in "cent" and "gem"; and "qu". Approximately 50 percent of the items contain consonant clusters and 50 percent do not.

Subtest 2 - Nonsense Word Production

These items do not represent meaningful words in English. The vowel sounds of the words are distributed to include approximately 50 percent short vowels, 25 percent long vowels and 25 percent special vowel sounds as described in the domain of Subtest 1. There are no restrictions on the consonant sounds that may be included.

Subtest 1 - Whole Word Production - Special Rules

This domain consists of words that contain the vowel sounds in the following words: partial, satiate, capture, join, saw, loud, food, good, term, corn, heart, bald, bulb. There is no restriction on the letters used to represent these sounds. Of course, other short and long vowel sounds may appear in words that have the sounds that are unique to this domain. There are no restrictions on the consonant sounds that may occur and words containing consonants that do not represent a sound (bomb) are included.

Subtest 8 - Nonsense Word Recognition

The items spoken by the examiner that are recognized by the examinee are identical to the items that are read orally by the examinee in Subtest 2. The items are in a different order than Subtest 2. This permits comparison of performance in the production and recognition modes. The four alternatives include: the correct item, an incorrect item with the first two letters the same as the standard, an incorrect item with the last two letters the same as the standard, and an incorrect item with the same letters as the standard in a new order.

Subtest 9 - Consonant Vowel Recognition

The items spoken by the examiner are the same as those of Subtest 5, except the sequence is different. The alternatives include: the correct item, an incorrect item with the first letter the same as the standard, an incorrect item with the second letter the same as the standard, and an incorrect item with neither letter the same as the standard.

Subtest 10 - Initial Letter Sound Recognition

The standards spoken by the examiner are 20 words beginning with a single consonant. There are four alternatives consisting of the correct letter, incorrect letters that are contained in the standard, and incorrect letters that are not in the standard.

III. Preliminary Technical Information

Reliability of the KIPT has not been properly and completely established. Additional data are needed. A suggestion that the reliability of the total test and individual reliabilities of the subtests may be adequate is provided in an article by one of the authors (Guthrie, 1973). Since the completion of the study, the KIPT has been revised and subtests in the two versions are not identical. Nonetheless, in this study the reliabilities of the subtests were:

1) word reading in context	.91
2) words flashed	.86
3) words unflashed	.88
4) nonsense word production	.86
5) long vowel production	.90
6) short vowel production	.90
7) consonant cluster production	.95
8) letter sound production	.86
9) letter naming	.22
10) nonsense word recognition	.83
11) consonant cluster recognition	.81
12) initial letter recognition	.88
13) final letter recognition	.83
14) auditory blending	.90
15) syllabication	.20
16) Total KIPT	.98

These reliabilities were computed with the KR21 formula with data from 38 Ss whose mean Gates-MacGinitie reading vocabulary was 2.14 grade level. See the article for further description of the subtests. It is important to note that the KR21 formula is intended for use with norm-referenced measures. A formula for computing the reliability of criterion-referenced tests was developed by Livingston (1972). That formula was not used for these computations since it was under development when the article was written. Livingston (1973) illustrates, however, that the criterion-referenced reliability is never lower and occasionally the reliabilities presented here are likely to be underestimates rather than overestimates of the true criterion-referenced reliability of the KIPT.

Validity of a criterion-referenced measure is likely to be more difficult to establish than validity of a typical norm-referenced measures. For example, a suitable criterion against which to judge a criterion-referenced measure may be difficult to identify. Preliminary indication of the validity of the KIPT was obtained from the correlation of .83 between the Nonsense Word Production Subtest of the KIPT and Gates-MacGinitie Vocabulary scores. Although the Letter Sound Production task correlated only .37 with Gates MacGinitie Vocabulary, Letter Sound Production correlated .53 with Consonant Cluster Production. The appropriate criterion for validating a given subtest may be a similar criterion-referenced test rather than a less similar norm-referenced measure. These issues demand additional thought and data.

Note of Caution:

Three reminders about the experimental version of the KIPT are in order. First, the authors do not make any claims about the validity or reliability of the measure until additional information is available. Second, interpretation drawn from the administration of the test may not be attributed to the authors. Third, the preliminary administrative manual does not attempt to meet the criteria of excellence described in the Standards for Educational and Psychological Tests and Manuals, (French & Michael, 1966).

Directions for Administration of the
Kennedy Institute Phonics Test

I. Materials

The examiner needs the following materials available for the administration of the test.

A. An examiner and examinee copy of subtests 1 - 10.

B. Pencils for use on subtests 8 - 10, and for examiner's use on entire test.

C. Paper or cardboard markers for examinees' use on test 8 - 10.

D. Scoring Sheet for recording examinees' responses.

E. Accuracy Profile Sheet for summarizing examinees' performance.

II. Seating arrangements and grouping.

A. Testing should be conducted in a relatively quiet room to prevent the examinee from being distracted and to enable the examiner to hear the oral responses of the examinee.

B. The examiner must test each examinee individually on subtests 1 - 7. Persons may be seated on opposite sides or adjacent to each other at a small table.

C. The examiner may administer subtests 8 - 10 in small groups. A group size of about 3 - 4 is suitable for children aged 5 - 8 years and for disabled readers; a group of 6 - 10 may be formed for normal children aged 9 and older.

III. Sequence and Time for Subtests

A. Subtests 1 - 10 should be administered in the order 1 - 10. The experience of the authors is that since subtests 1-7 are ordered in decreasing difficulty, examinees are encouraged by success as they progress through the subtests and their motivation to perform well is sustained. In addition, subtests 8 - 10 must follow 1 - 7 since the items on 8 and 9 are the same as the items on subtests 2 and 5. Since the former requires oral production and the latter requires recognition, it is likely that administration of 8 and 9 before 2 and 5 may inflate performance on 2 and 5 although the reverse is not likely.

B. In general, the time required to administer subtests 1 - 7 is 15 to 30 minutes and the time to administer subtests 8 - 10 is 10 to 15 minutes. The total time varies from 25 to 45 minutes depending on the age and reading level of the examinee. Young children (aged 5-8) and disabled readers (grade level 1.0-2.5) take more time to complete the test than older students and more advanced readers. Subtests 1 - 10 may be given to an individual child in one or two sessions. Subtests 1 - 7 may be given in one or two sessions as needed and subtests 8 - 10 should be given in one session.

IV. Administration of Subtests

The four alternate forms may be used interchangeably and are approximately equal in difficulty level. Subtests from one form should not be mixed with subtests from other forms. Administration of subtests 1 - 7 is conducted individually. Instructions for each subtest are provided on the examiner's copy of the subtest. These instructions should be read aloud to the child. All of the subtests have examples. After hearing the instructions the examinee shoul[illegible]form the examples. The examiner may provide any help that is needed to enable the child to perform the examples correctly.

Additional examples should not be given and meanings of the words should not be emphasized.

When the examinee is working on the test items, no assistance should be given in pronouncing the words. If an examinee hesitates on one item for a long time or provides a partial response such as the sound of the first letter in a short vowel word, encouragement may be given by saying, "are you going to try that one" or "what is your answer on that one." Encourage the examinee to attempt the next item if he cannot perform an item and perseverates on it. If the examinee produces the separate sounds for an item by saying "b"...."i" for "bi", the examiner may suggest that the examinee "put the sounds together". Lengthy assistance should not be given and responses which are not adequately blended are scored as incorrect. Do not provide specific feedback about whether he is right or wrong on any items. The examinee may use a marker or pointer or the examiner may assist him in maintaining his place on the testing page. It is important that a given subtest be discontinued if the examinee makes six consecutive errors. The next subtest should be started immediately after such a discontinuation.

Administration of subtests 8 - 10 may be conducted in groups as described previously. Directions for each subtest are provided on the examiner's copy and should be read to the examinee. The examinees should perform the examples and assistance may be given to insure that the examples are answered correctly and that the examinees understand the task. Examinees should be encouraged to use markers to maintain their place on the page. The examiner may pronounce each stimulus word two times, although he should not return to an item that was previously administered to assist examinees who made an error. All 20 items on each of the subtests 8 - 10 should be administered to all examinees.

V. Guidelines for Judging and Scoring Performance

A. Subtest 1 - Whole Word Production - Special Rules

Score as correct the standard pronunciation of the word. Make minimal allowance for dialect and accent differences. If an examinee has a non-standard black dialect or a marked Bostonian accent, for example, accept oral responses that are consistent with his linguistic system.

B. Subtest 2 - Nonsense Word Production

Score as correct responses that follow standard phonetic rules. Although a complete system of phonetic rules cannot be provided here, the following guidelines will suffice for judging responses to these items. If the word contains one vowel, either the short or long sound may be heard, e.g. "rax" should rhyme with "tax" and "tright" should rhyme with "might". If a word contains two vowels, the vowels may be pronounced with any of the sounds which those vowels may represent in English. For instance, "goot" may be correctly pronounced as rhyming with either "foot" or "boot". Vowels that occur in the beginning of a word may be pronounced with either the short or long sound. For instance, the "o" in "ont" may be pronounced as the "o" in "object" or "only". The diacritical marks used by the examiner in Subtest 8 should not be used in judging the adequacy of responses on this subtest.

C. Subtest 3 - Long Vowel Word Production

Score as correct the standard pronunciation of the word with minimal allowance for dialect or accent differences.

D. Subtest 4 - Short Vowel Word Production

Score as correct the standard pronunciation of the word with minimal allowance for dialect differences.

E. Subtest 5 - Consonant Vowel Production

Score as correct the standard consonant sound with allowances for different consonant sounds for certain letters. For example, in "ge" the "g" may be pronounced as the "g" in "get" or "gesture". For vowels, either the long or short sound is acceptable. For instance, in "da" the "a" may be sounded as in "dam" or "dame".

F. Subtest 6 - Letter Sound Production

Score as correct any of the sounds that a consonant may make in the initial position of a word. Note that "c" may be hard or soft as in "cat" or "city".

G. Subtest 7 - Letter Naming

Score as correct letter names such as "bee" for "b".

H. Subtest 8 - Nonsense Word Recognition

Score as correct the standard spoken by the examiner if the examinee circles it correctly among the four alternatives.

I. Subtest 9 - Consonant Vowel Recognition

Score as correct the standard spoken by the examiner if the examinee circles it correctly among the four alternatives.

J Subtest 10 - Initial Letter Sound Recognition

Score as correct the initial letter of the word spoken by the examiner if the examinee circles it.

VI. Constructing Examinee Profile

While the examinee is performing Subtests 1 - 7, the examiner records each response as correct or incorrect on the Scoring Sheet. Performance on Subtests 8 - 10 is recorded on the sheet when the examinees complete the section. The total number correct for each subtest is then computed and recorded at the bottom of the Scoring Sheet. Information from the Scoring Sheet is then placed on the Profile Sheet. The number of correct responses on each subtest is entered on the Profile Sheet in the column labeled "number correct". The percent correct, calculated as number correct divided by total possible correct, is entered in the section entitled "percent correct". This information may be entered as a dot in the appropriate row. The Profile is created by connecting the points with straight lines.

REFERENCES

French, J.W. and Michael, W.B. Standards for Educational and Psychological Tests and Manuals. Washington, D.C. American Psychological Association, 1966.

Glaser, R. Instructional technology and the measurement of learning outcomes. American Psychologist, 1963, 18, 519-521.

Guthrie, J.T. Models of reading and reading disability. Journal of Educational Psychology, 1973, 65 (1), 9-18.

Hambleton, R.K. and Novick, M.R. Toward an integration of theory and method for criterion-referenced test. ACT Research Report No. 53. Iowa City, American College Testing Program, 1972.

Kriewall, Thomas E. Aspects and applications of criterion-referenced tests. Technical Paper No. 103 Institute for Educational Research, Downers Grove, Illinois, 1972.

Livingston, S.A. Classical test theory for criterion-referenced tests. Dissertation submitted to Johns Hopkins University, Baltimore, 1973.

Livingston, S.A. Criterion-referenced applications of classical test theory. Journal of Educational Measurement, 1972, 9, 13-26.

Lyman, H.B. Talking test scores: Criterion-referenced testing. NCME Measurement News, 1972, 15 (3), 3.

Nitko, Anthony J. A model for criterion-referenced tests based on use. Paper presented at the annual meeting of the American Educational Research Association. New York, February, 1971.

Popham, W.J. (ed.) Criterion-Referenced Measurement: An Introduction. Englewood Cliffs, New Jersey. Educational Technology Publications, 1971.

Roudabush, G.E. and Green, D.R. Some reliability problems in a criterion-referenced test. Paper presented at the annual meeting of the American Educational Research Association Meeting, New York, February, 1971.

Wentling, Tim L. Mastery versus nonmastery instruction with varying test item feedback treatments. Journal of Educational Psychology, 1973, 65 (1), 50-58.

BEST COPY AVAILABLE

PRODUCTION | RECOGNITION

ubtest	1	1	2	3	4	5	6	7		8	9	10
1	–	21	+	–	+	+	+	+		+	+	+
2	–	22	–	–	–	+	+	+		+	+	+
3	–	23	–	–	+	+	+	+		+	+	+
4	–	24	+	–	–	+	+	+		+	+	+
5	–	25	–	–	+	–	+	+		–	+	+
6	+	26	–	+	–	+	+	–		+	+	+
7	–	27	+	–	–	+	+	+		+	+	+
8	–	28	–	–	+	+	+	+		+	+	+
9	+	29	+	+	–	+	+	+		+	+	+
10	–	30	–	–	+	–	+	+		+	+	+
11	–		–	–	–	–	–	+		+	+	+
12	–		–	–	+	+	+	–		+	+	+
13	–		–	–	–	+	+	+		+	+	+
14	–		+	–	–	+	+	+		+	+	+
15	–		–	–	–	–	+	+		+	+	+
16	stop		–	stop	–	+	+	+		–	+	+
17			–		–	+	+	+		+	+	+
18			+		–	–	+	+		+	+	+
19			–		stop	–	+	+		+	+	+
20			–			+	–	+		+	+	+
orrect		2	6	2	6	14	18	18		18	20	[illegible]

Name John Smith

Date 9-6-73 Form F Age 11-2

Examiner [illegible]

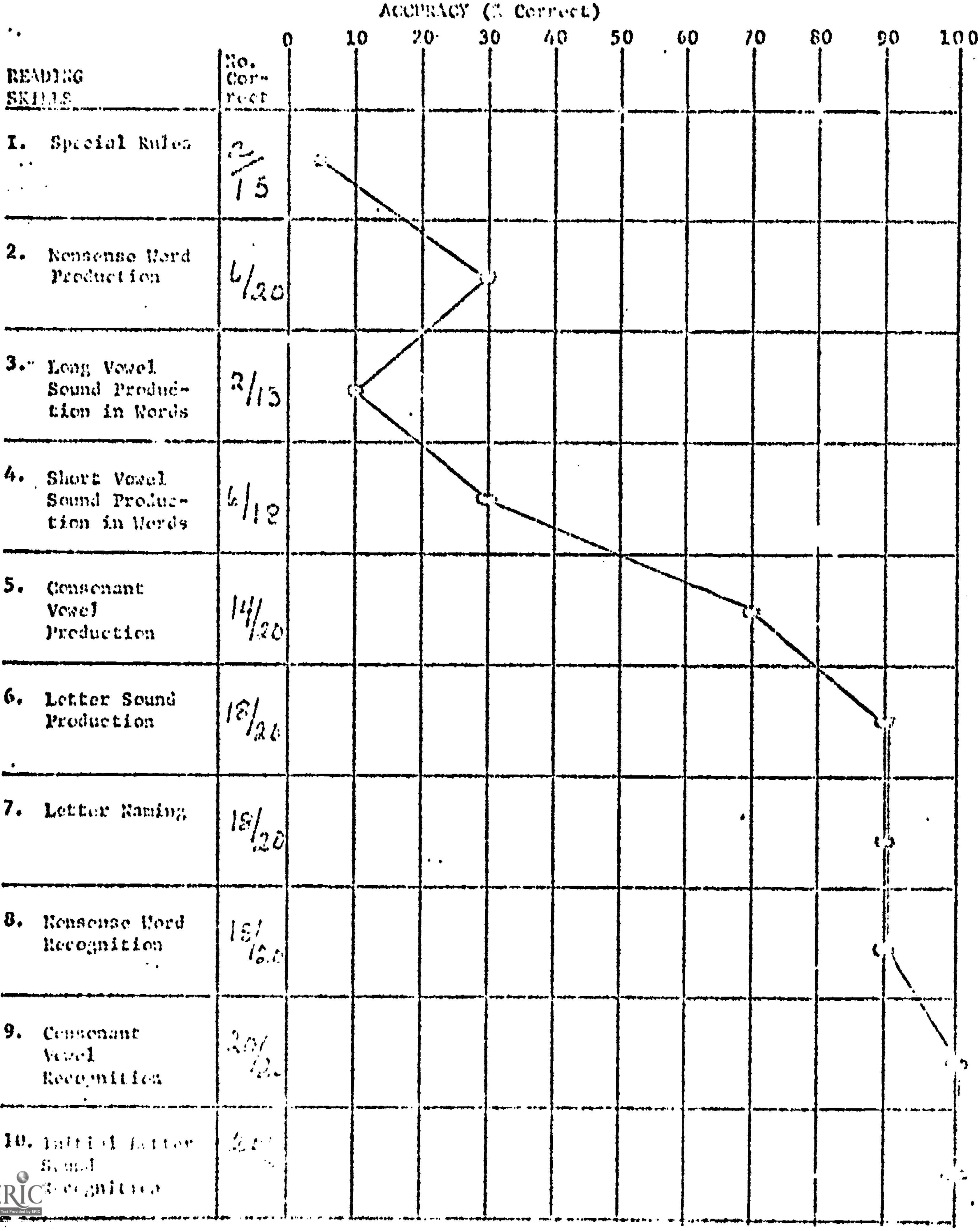
KIPT PROFILE

BEST COPY AVAILABLE

NAME John Smith FORM F

DATE OF EXAMINATION 9-5-73 GRADE PLACEMENT ungraded

EXAMINER M. Seifert SCHOOL Kennedy School



TEST 1 -- WHOLE WORD PRODUCTION -- SPECIAL RULES

A. car

B. moon

1. cried
2. kneel
3. thick
4. shirt
5. silly
6. noise
7. storm
8. chrome
9. extend
10. earth
11. nature
12. wring
13. judge
14. cease
15. bloom
16. favor
17. board
18. hearth
19. column
20. needle
21. debris
22. might
23. yawn
24. growl
25. heard
26. chore
27. bald
28. chalk
29. dye
30. cause

Directions: I would like you to read as many of these words as you can. Let's try the practice ones. (Examiner pronounce examples correctly if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

TEST 1 -- WHOLE WORD PRODUCTION --
SPECIAL RULES

A. car

B. moon

1. cried
2. kneel
3. thick
4. shirt
5. silly
6. noise
7. storm
8. chromo
9. extend
10. earth
11. nature
12. wring
13. judge
14. cease
15. bloom
16. favor
17. board
18. hearth
19. column
20. needle
21. debris
22. might
23. yawn
24. growl
25. heard
26. chore
27. bald
28. chalk
29. dye
30. cause

TEST 2 -- NONSENSE WORD PRODUCTION

A. fim

B. ack

1.	gak	11.	greer
2.	spro	12.	teefer
3.	ert	13.	ipsem
4.	grud	14.	glore
5.	freel	15.	shuff
6.	poit	16.	rocress
7.	clid	17.	staum
8.	tave	18.	thulk
9.	ock	19.	marbop
10.	scune	20.	endoke

Directions: This page has all nonsense words. That means that none of the words make any sense. They are just letter sounds put together. Lets try the practice ones. (Examiner pronounce the practice words correctly if child does not; then have child repeat examples.) Discontinue subtest after 6 consecutive errors.

C

TEST 2 -- NONSENSE WORD PRODUCTION

A. fim

B. ack

1. gak
2. spro
3. ert
4. grud
5. freel
6. poit
7. clid
8. tave
9. ock
10. scune
11. greer
12. teefer
13. ipsem
14. glore
15. shuff
16. rocress
17. staum
18. thulk
19. marbop
20. endoke

C

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS

A. make

B. free

1.	rail	11.	prey
2.	heap	12.	bolt
3.	lime	13.	tune
4.	robe	14.	cheat
5.	crude	15.	clue
6.	loaf	16.	great
7.	try	17.	mile
8.	cage	18.	slow
9.	queen	19.	fear
10.	blind	20.	fuse

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce the practice words correctly if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

C

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS

A. make

B. free

1.	rail	11.	prey
2.	heap	12.	bolt
3.	lime	13.	tune
4.	robe	14.	cheat
5.	crude	15.	clue
6.	loaf	16.	great
7.	try	17.	mile
8.	cage	18.	slow
9.	queen	19.	fear
10.	blind	20.	fuse

C

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. sit

B. man

1.	sad	11.	rung
2.	fed	12.	swap
3.	sip	13.	lit
4.	nod	14.	clamp
5.	young	15.	bun
6.	snap	16.	ham
7.	web	17.	next
8.	fist	18.	slot
9.	bread	19.	clip
10.	log	20.	gum

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce practice words correctly if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

C

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. sit

B. man

1. sad
2. fed
3. sip
4. nod
5. young
6. snap
7. web
8. fist
9. bread
10. log
11. rung
12. swap
13. lit
14. clamp
15. bun
16. ham
17. next
18. slot
19. clip
20. gum

C

TEST 5 -- PRODUCTION OF VOWEL-
CONSONANT COMBINATIONS

A. ep

B. co

1. eg
2. ca
3. ek
4. zu
5. ag
6. on
7. ni
8. vi
9. af
10 la
11. he
12. os
13. te
14. mo
15. tu
16. in
17. az
18. oh
19. re
20. uh

Directions: This page has two-letter items. Some make sense and some are nonsense words. Tell me what the letters say together. Lets try the practice ones. (The child may use either the long or short vowel sound. If the child cannot pronounce the examples using either sound, the examiner should pronounce them using a short vowel sound; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

C

TEST 5 -- PRODUCTION OF VOWEL - CONSONANT COMBINATIONS

A. ep

B. co

1.	eg	11.	hc
2.	ca	12.	os
3.	ek	13.	te
4.	zu	14.	mo
5.	ag	15.	tu
6.	on	16.	in
7.	ni	17.	az
8.	vi	18.	oh
9.	af	19.	re
10.	la	20.	uh.

C

TEST 6 -- LETTER SOUND PRODUCTION

A. h

B. t

o	c	g	f	s	y	e	v	a	w
d	i	u	z	p	m	j	n	l	b

Directions: I want you to tell me the sound each one of these letters make. Lets try the practice ones. (Examiner pronounce examples if the child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

* *

TEST 7 -- LETTER NAMING

A. s

B. v

e	a	z	q	w	d	r	f	m	x
y	n	k	n	l	c	p	o	i	b

Directions: Tell me the names of each of these letters. Lets try the practice ones. (Examiner pronounce examples if the child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

TEST 6 -- LETTER SOUND PRODUCTION

A. h

B. t

o c g f s y e v a w

d i u z p m j n l b

****

TEST 7 -- LETTER NAMING

A. s

B. v

u a z q w d r f m x

y n k n l c p o i b

C

TEST 8 -- RECOGNITION OF THE VISUAL FORM OF NONSENSE WORDS

A. fĭm

B. ăck

1. ŏck
2. clĭd
3. tāve
4. scūne
5. poit (boil)
6. frēel
7. găk
8. ert (germ)
9. grŭd
10. sprō
11. thŭlk
12. ĕndōke
13. staum (audit)
14. rŏcrĕss
15. marbop (star)
16. tēefer
17. ĭpsĕm
18. greer (sneer)
19. shŭff
20. glore (or)

Directions: This page has rows of all nonsense words. Each row has four words. I will say a nonsense word and your job is to look across the row and find the same word I said. Put a circle around the word. Lets try the practice ones. Each child completes items 1 - 20.

C

TEST 8 -- RECOGNITION OF THE VISUAL FORM OF NONSENSE WORDS

A.	lim	fim	mif	fir
B.	eck	cak	ack	acf
1.	uck	cko	ock	ocr
2.	clid	frid	dilc	clon
3.	feve	taro	vate	tave
4.	unsce	scoru	nurne	scune
5.	poti	poit	pouf	jait
6.	froof	freol	thoel	elfer
7.	fak	gar	gak	kag
8.	ert	ret	erm	ort
9.	grud	jaud	rugd	grel
10.	pros	spne	noro	spro
11.	thulk	kulth	thoff	flelk
12.	entuhe	endoke	dokeen	urhake
13.	tausm	stown	staum	dreum
14.	sonnass	resscor	rocress	ronoarn
15.	mastug	marbop	nontop	opmarb
16.	ofeort	tenlon	fooher	teefer
17.	ipnow	semip	ipsem	ognem
18.	phoer	greer	groow	erger
19.	sholl	uffsh	shuff	nooff
20.	glore	glorn	oregl	phure

C

TEST 9 -- RECOGNITION OF VOWEL-CONSONANT COMBINATIONS

A. ep

B. co

1.	ni	11.	az
2.	af	12.	re
3.	on	13.	oh
4.	vi	14.	in
5.	la	15.	uh
6.	zu	16.	tu
7.	ca	17.	he
8.	eg	18.	te
9.	ek	19.	os
10.	ag	20.	mo

Directions: This page has rows of consonant vowel combinations. I will say the sound that two letters make together and you put a circle around the two letters that make that sound. Lets try the practice ones. (Examiner pronounce all combinations with the short vowel sound.) Each child completes items 1 - 20.

C

TEST 9 -- RECOGNITION OF VOWEL-CONSONANT COMBINATIONS

A.	eb	op	ep	ah
B.	cu	co	da	mo
1.	no	ni	ca	gi
2.	aw	du	af	ef
3.	od	an	on	su
4.	va	pi	ko	vi
5.	la	ta	lu	ge
6.	zu	ri	za	tu
7.	co	ca	ne	da
8.	ag	em	ov	eg
9.	uk	eh	iz	ek
10.	ap	og	ag	ik
11.	ez	az	od	al
12.	ru	re	ke	sa
13.	oh	ih	ow	uc
14.	ih	af	in	en
15.	um	ih	av	uh
16.	ta	tu	se	fu
17.	he	li	hu	ke
18.	te	fe	tu	po
19.	ot	es	os	ig
20.	ma	vo	ne	mo

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

A. salamander

B. numerical

1. luscious
2. harmony
3. zither
4. pendulum
5. temporary
6. young
7. casual
8. gumption
9. kangaroo
10. notorious
11. symphony
12. willow
13. bovine
14. furious
15. judicial
16. marvelous
17. rascal
18. vigorous
19. dynamite
20. natural

Directions: Each row on this page has four letters. I will say a word. You look at the letters in the row and put a circle around the letter that you hear at the beginning of the word. Lets try the practice ones. Each child completes items 1 - 20.

C

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

A.	l	m	d	s
B.	m	n	r	l
1.	s	l	v	c
2.	r	y	h	m
3.	t	r	z	h
4.	m	d	n	p
5.	t	m	r	p
6.	y	n	g	h
7.	c	l	z	s
8.	m	g	n	p
9.	g	r	k	n
10.	t	s	n	r
11.	h	p	m	s
12.	l	v	s	w
13.	v	n	b	d
14.	f	s	s	r
15.	d	c	j	l
16.	r	m	v	l
17.	s	c	r	l
18.	g	v	s	r
19.	d	t	n	m
20.	t	r	l	n

TEST 1 -- WHOLE WORD PRODUCTION
SPECIAL RULES

A. boy

B. down

1. calves
2. sign
3. birth
4. term
5. have
6. gas
7. straw
8. expect
9. rather
10. ghost
11. tension
12. wreck
13. listen
14. lye
15. stood
16. dollar
17. charm
18. gulp
19. autumn
20. scale
21. whole
22. hour
23. point
24. pearl
25. spurt
26. court
27. sugar
28. hedge
29. prose
30. crown

Directions: I would like you to read as many of these words as you can. Let's try the practice ones. (Examiner pronounce examples if child does not; then have child repeat the examples.) Discontinue the subtest after six consecutive errors.

D

TEST 1 -- WHOLE WORD PRODUCTION
SPECIAL RULES

A. boy

B. down

1. calves
2. sign
3. birth
4. term
5. have
6. gas
7. straw
8. expect
9. rather
10. ghost
11. tension
12. wreck
13. listen
14. lye
15. stood
16. dollar
17. charm
18. gulp
19. autumn
20. scale
21. whole
22. hour
23. point
24. pearl
25. spurt
26. court
27. sugar
28. hedge
29. prose
30. crown

D

TEST 2 -- NONSENSE WORD PRODUCTION

A. kag

B. uck

1. paf
2. smay
3. floy
4. elch
5. drute
6. urb
7. thab
8. hice
9. fing
10. flome
11. treave
12. ermy
13. upple
14. clow
15. spack
16. dectow
17. stee
18. nedge
19. magsto
20. stipap

Directions: This page has all nonsense words. That means that none of the words make any sense. They are just letter sounds put together. Lets try the practice ones. (Examiner correctly pronounce the practice words if the child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

D

TEST 2 -- NONSENSE WORD PRODUCTION

A. kag

B. uck

1. paf
2. smay
3. floy
4. elch
5. drute
6. urb
7. thab
8. hice
9. fing
10. flome
11. treave
12. ermy
13. upple
14. clow
15. spack
16. dectow
17. stee
18. nedge
19. magsto
20. stipap

D

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS

A. home

B. tail

1. pain
2. steep
3. pipe
4. most
5. mute
6. note
7. grind
8. shake
9. peak
10. wild
11. freight
12. cope
13. chew
14. cream
15. bruise
16. pare
17. tire
18. roll
19. fierce
20. tube

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce the practice words correctly if the child does not; then have child repeat the examples.) Discontinue the subtest after six consecutive errors.

D

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS

A. home

B. tail

1. pain
2. steep
3. pipe
4. most
5. mute
6. note
7. grind
8. shake
9. peak
10. wild
11. freight
12. cope
13. chew
14. cream
15. bruise
16. pare
17. tire
18. roll
19. fierce
20. tube

D

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. bat

B. pen

1. jam
2. net
3. him
4. cot
5. ton
6. grab
7. sped
8. swim
9. bent
10. hog
11. spun
12. swan
13. mix
14. smack
15. rug
16. nag
17. head
18. trot
19. miss
20. hunt

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce practice words correctly if child does not; then have the child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

D

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. bat

B. pen

1. jam
2. net
3. him
4. cot
5. ton
6. grab
7. sped
8. swim
9. bent
10. hog
11. spun
12. swan
13. mix
14. smack
15. rug
16. nag
17. head
18. trot
19. miss
20. hunt

D

TEST 5 -- PRODUCTION OF VOWEL-
CONSONANT COMBINATIONS

A. ic

B. ho

1.	ap	11.	ol
2.	di	12.	ib
3.	ji	13.	ze
4.	we	14.	im
5.	ba	15.	hi
6.	es	16.	av
7.	je	17.	wo
8.	et	18.	op
9.	af	19.	iz
10.	wi	20.	ku

Directions: This page has two-letter items. Some make sense and some are nonsense words. Tell me what the letters say together. Lets try the practice ones. (Examiner, the child may use either the long or short vowel sound. If child cannot pronounce the examples using either sound, the examiner should pronounce them using a short vowel sound; then have child repeat the examples.) Discontinue the subtest after six consecutive errors.

D

TEST 5 -- PRODUCTION OF VOWEL - CONSONANT COMBINATIONS

A. ic

B. ho

1. ap
2. di
3. ji
4. we
5. ba
6. es
7. je
8. et
9. af
10. wi
11. ol
12. ib
13. ze
14. im
15. hi
16. av
17. wo
18. op
19. iz
20. ku

D

TEST 6 -- LETTER SOUND PRODUCTION

A. s

B. m

u	e	b	r	a	h	w	z	f	n
y	d	l	k	t	g	j	i	p	v

Directions: I want you to tell me the sound each one of these letters make. Lets try the practice ones. (Examiner pronounce examples if the child does not; then have child repeat the examples.) Discontinue the subtest after six consecutive errors.

* *

TEST 7 -- LETTER NAMING

A. d

B. g

e	h	b	l	v	u	c	n	m	y
a	o	x	p	r	f	z	q	k	t

Directions: Tell me the names of each of these letters. Lets try the practice ones. (Examiner pronounce examples if the child does not; then have child repeat the examples.) Discontinue the subtest after six consecutive errors.

D

TEST 6 -- LETTER - SOUND PRODUCTION

A. s

B. m

u	c	b	r	a	h	w	z	f	n
y	d	l	k	t	g	j	i	p	v

****

TEST 7 -- LETTER NAMING

A. d

B. g

e	h	b	l	v	u	c	n	m	y
a	o	x	p	r	f	z	q	k	t

D

TEST 8 -- RECOGNITION OF THE VISUAL FORM OF NONSENSE WORDS

A. kăg

B ŭck

BEST COPY AVAILABLE

1. thŭb
2. fĭng
3. urb (slur)
4. hīce
5. flōme
6. drūte
7. smāy
8. păf
9. floy (boy)
10. elch (felt)
11. stēe
12. măgstō
13. dĕctow (cow)
14. nedge (hedge)
15. stīpăp
16. spăck
17. ŭpple
18. ermy (germ)
19. trēave
20. clow (cow)

Directions: This page has rows of all nonsense words. Each row has four words. I will say a nonsense word and your job is to look across the row and find the same word I said. Put a circle around the word. Lets try the practice ones. Each child completes items 1 - 20.

D

TEST 8 -- RECOGNITION OF THE VISUAL FORM OF NONSENSE WORDS

A.	hag	kaf	kag	gak
B.	ack	cku	ucr	uck
1.	baht	thab	thom	srab
2.	lang	fiot	fing	ginf
3.	urw	urb	arb	rbu
4.	hice	cihe	loce	hito
5.	fluwt	mofle	flome	trame
6.	droly	utder	snate	drute
7.	smen	toay	ausm	smay
8.	paf	caf	fap	paw
9.	flar	floy	oyfl	shoy
10.	mach	elch	chle	elpt
11.	lnce	sete	stee	stuf
12.	magsto	mapron	stomag	unjrto
13.	dectow	towdec	bonfow	deuhon
14.	denge	nedge	nelor	ufage
15.	flojap	strome	papsti	stipap
16.	ackps	spoth	spack	nouck
17.	uphan	upple	puple	sonnle
18.	ermy	erui	mery	nomy
19.	eatrev	trouny	hniove	treave
20.	frow	wocl	clow	clum

TEST 9 -- RECOGNITION OF VOWEL-
CONSONANT COMBINATIONS

A. ic

B. ho

1. wi	11. wo
2. et	12. iz
3. je	13. ku
4. af	14. av
5. es	15. op
6. ba	16. ib
7. ji	17. ze
8. ap	18. hi
9. di	19. ol
10. we	20. im

Directions: This page has rows of consonant vowel combinations. I will say the sound that two letters make together and you put a circle around the two letters that make that sound. Lets try the practice ones. (Examiner pronounce all combinations with the short vowel sound.) Each child completes items 1 - 20.

D

TEST 9 -- RECOGNITION OF VOWEL-CONSONANT COMBINATIONS

A.	uc	ic	im	ju
B.	mo	de	ho	hi
1.	wa	du	wi	ci
2.	eh	av	et	ot
3.	je	ja	ne	wi
4.	af	ef	ig	ar
5.	ew	es	us	ov
6.	bo	ta	su	ba
7.	fi	jo	ji	pu
8.	az	ap	op	gi
9.	di	lu	xi	da
10.	wu	bo	de	we
11.	wa	wo	ru	no
12.	ik	az	iz	ef
13.	ku	mu	ke	bi
14.	lo	av	aw	iv
15.	av	ep	op	ob
16.	ig	ab	oz	ib
17.	za	se	mi	ze
18.	gu	hi	ha	ni
19.	ol	el	of	da
20.	oh	um	im	in

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

A. personal

B. valuable

1. demise
2. conquest
3. buskin
4. zygote
5. hatchet
6. sanitary
7. reference
8. landscape
9. jealous
10. victorious
11. population
12. musician
13. torment
14. nebulous
15. fantastic
16. wealth
17. yolk
18. geography
19. kayak
20. gauntlet

Directions: Each row on this page has four letters. I will say a word. You look at the letters in the row and put a circle around the letter that you hear at the beginning of the word. Lets try the practice ones. (Examiner say each word twice.) Each child completes items 1 - 20.

D

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

A.	r	s	n	p
B.	n	b	v	l
1.	m	z	d	s
2.	s	c	n	t
3.	k	b	s	n
4.	y	g	z	t
5.	c	t	l	h
6.	n	s	r	t
7.	r	f	c	n
8.	s	p	n	l
9.	j	l	w	s
10.	c	v	s	t
11.	l	t	p	n
12.	s	n	c	m
13.	m	t	r	n
14.	s	l	n	b
15.	f	n	t	s
16.	l	w	h	t
17.	l	w	k	y
18.	r	g	p	h
19.	w	k	b	y
20.	l	t	g	n

TEST 1 -- WHOLE WORD PRODUCTION
SPECIAL RULES

A. chin

B. girl

1. hiking
2. gnaw
3. cable
4. toy
5. couch
6. tore
7. starch
8. thank
9. chunk
10. phone
11. special
12. thumb
13. castle
14. praise
15. cent
16. worn
17. floor
18. heart
19. knife
20. scope
21. whom
22. half
23. true
24. swoop
25. nerve
26. pour
27. sure
28. taught
29. letter
30. exit

Directions: I would like you to read as many of these words as you can. Lets try the practice ones. (Examiner pronounce the examples if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

E

TEST 1 -- WHOLE WORD PRODUCTION
SPECIAL RULES

A. chin

B. girl

1. hiking
2. gnaw
3. cable
4. toy
5. couch
6. tore
7. starch
8. thank
9. chunk
10. phone
11. special
12. thumb
13. castle
14. praise
15. cent
16. worn
17. floor
18. heart
19. knife
20. scope
21. whom
22. half
23. true
24. swoop
25. nerve
26. pour
27. sure
28. taught
29. letter
30. exit

E

TEST 2 -- NONSENSE WORD PRODUCTION

A. kig

B. emp

1.	rops	11.	woto
2.	spe	12.	siness
3.	smorf	13.	etting
4.	sten	14.	douf
5.	plake	15.	tront
6.	ghaw	16.	kenner
7.	ance	17.	blumb
8.	nire	18.	wreps
9.	eld	19.	iby
10.	thew	20.	creteck

Directions: This page has all nonsense words. That means that none of the words make any sense. They are just letter sounds put together. Lets try the practice ones. (Examiner correctly pronounce the practice words if the child does not.) Then have the child repeat the examples. Discontinue the subtest after 6 consecutive errors.

E

TEST 2 -- NONSENSE WORD PRODUCTION

A. kig

B. emp

1. rops
2. spe
3. smorf
4. ston
5. plake
6. ghaw
7. ance
8. nire
9. eld
10. thew
11. wote
12. siness
13. etting
14. douf
15. tront
16. kenner
17. blumb
18. wreps
19. iby
20. creteck

E

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS.

A. goat

B. cake

1.	lace	11.	vein
2.	seem	12.	soul
3.	pine	13.	glue
4.	role	14.	beak
5.	cube	15.	grew
6.	quote	16.	scare
7.	mild	17.	tribe
8.	braid	18.	juice
9.	high	19.	steer
10.	plead	20.	crow

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce the practice words correctly if the child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

E

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS

A. goat

B. cake

1. lace
2. seem
3. pine
4. role
5. cube
6. quote
7. mild
8. braid
9. high
10. plead
11. vein
12. soul
13. glue
14. beak
15. grew
16. scare
17. tribe
18. juice
19. steer
20. crow

E

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. pet

B. bad

1.	nap	11.	plum
2.	fed	12.	wasp
3.	tip	13.	fit
4.	fog	14.	damp
5.	done	15.	rub
6.	flat	16.	tax
7.	wet	17.	spread
8.	slim	18.	song
9.	rest	19.	limp
10.	hop	20.	sunk

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce practice words correctly if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

E

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. pet

B. bad

1.	nap	11.	plum
2.	fed	12.	wasp
3.	tip	13.	fit
4.	fog	14.	damp
5.	done	15.	rub
6.	flat	16.	tax
7.	wet	17.	spread
8.	slim	18.	song
9.	rest	19.	limp
10.	hop	20.	sunk

L

TEST 5 -- PRODUCTION OF VOWEL - CONSONANT COMBINATIONS

A. ok

B. ri

1.	ta	11.	hu
2.	do	12.	si
3.	je	13.	mu
4.	ha	14.	om
5.	pi	15.	ho
6.	da	16.	ic
7.	il	17.	ub
8.	ru	18.	ga
9.	ti	19.	po
10.	du	20.	et

Directions: This page has two-letter items. Some make sense and some are nonsense words. Tell me what the letters say together. Lets try the practice ones. (The child may use either the long or short vowel sound. If the child cannot pronounce the examples using either sound, the examiner should pronounce them using a short vowel sound; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

TEST 5 -- PRODUCTION OF VOWEL - CONSONANT COMBINATIONS

A. ok

B. ri

1. ta
2. do
3. je
4. ha
5. pi
6. da
7. il
8. ru
9. ti
10. du
11. hu
12. si
13. mu
14. om
15. ho
16. ic
17. ub
18. ga
19. po
20. et

E

TEST 6 -- LETTER SOUND PRODUCTION

A. l

B. d

r	n	w	e	z	v	y	p	h	g
i	m	f	t	b	c	o	s	j	u

Directions: I want you to tell me the sound each one of these letters make. Lets try the practice ones. (Examiner pronounce the examples if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

* *

TEST 7 -- LETTER NAMING

A. b

B. k

q	z	u	a	f	v	i	d	j	c
x	w	n	m	e	r	h	g	o	s

Directions: Tell me the names of each of these letters. Lets try the practice ones. (Examiner pronounce the examples if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

TEST 6 -- LETTER SOUND PRODUCTION

A. l

B. d

r n w e z v y p h g

i m f t b c o s j u

* *

TEST 7 -- LETTER NAMING

A. b

B. k

q z u a f v i d j c

x w n m e r h g o s

E

TEST 8 -- RECOGNITION OF THE VISUAL FORM
OF NONSENSE WORDS

A. kig

B. emp

1. old (felt)
2. ance
3. tnew (blew)
4. nire
5. ghaw (saw)
6. spe
7. smorf (or)
8. plake
9. rops
10. sten
11. wreps
12. creteck
13. blumb
14. kenner
15. iby
16. douf (out)
17. etting
18. tront
19. wote
20. siness

Directions: This page has rows of all nonsense words. Each row has four words. I will say a nonsense word and your job is to look across the row and find the same word I said. Put a circle around the word. Lets try the practice ones. Each child completes items 1 - 20.

E

TEST 8 -- RECOGNITION OF THE VISUAL FORM OF NONSENSE WORDS

A.	tig	kig	gik	kif
B.	omp	emf	pem	emp
1.	elh	dle	eld	ald
2.	ance	omce	nace	antu
3.	thou	hewt	frew	thew
4.	sare	nimo	rine	nire
5.	ghaw	ghun	awhg	slaw
6.	upe	spe	esp	spa
7.	gnirf	rofms	smorf	smaut
8.	plohn	truke	plake	kaple
9.	rops	pors	nups	roge
10.	flen	sten	stau	enst
11.	stups	perws	wrojr	wreps
12.	teckere	snohock	creteck	cruholh
13.	blont	umlbe	shomb	blumb
14.	kenner	kewrsy	houmer	nenerk
15.	byi	iby	ibt	oby
16.	foud	dont	douf	lauf
17.	etlome	etting	allong	tinget
18.	tront	trawr	onttr	snunt
19.	towe	nuto	wote	wofn
20.	sihoan	essnis	hutoss	sinoss

TEST 9 -- RECOGNITION OF CONSONANT VOWEL COMBINATIONS

A. ri

B. ok

1.	il	11.	po
2.	ti	12.	ic
3.	du	13.	ga
4.	ru	14.	ub
5.	da	15.	et
6.	pi		om
7.	je	17.	si
8.	ta	18.	mu
9.	do	19.	hu
10.	ha	20.	ho

Directions: This page has rows of consonant vowel combinations. I will say the sound that two letters make together and you put a circle around the two letters that make that sound. Lets try the practice ones. (Examiner pronounce all combinations with the short vowel sound.) Each child completes items 1 - 20.

TEST 9 -- RECOGNITION OF CONSONANT VOWEL COMBINATIONS

BEST COPY AVAILABLE

A.	ni	fu	ti	ra
B.	el	ek	iv	ak
1.	af	il	ol	it
2.	ti	ta	fi	de
3.	da	du	zo	bu
4.	ra	di	zu	ru
5.	gu	do	da	ba
6.	pi	sa	pu	bi
7.	je	ju	ho	re
8.	yu	fa	ta	to
9.	de	bo	mi	do
10.	he	ju	ha	ba
11.	pi	po	lu	bo
12.	id	ic	ze	ac
13.	gi	pa	co	ga
14.	ub	af	ud	ob
15.	ef	uk	et	ot
16.	bi	om	em	ov
17.	ri	ca	so	si
18.	nu	mi	mu	jo
19.	ha	nu	hu	re
20.	ho	mo	hi	tu

E

BEST COPY AVAILABLE

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

A. context

B. manager

1. nasturium
2. function
3. walnut
4. yesterday
5. gallivant
6. kerosene
7. haunch
8. taciturn
9. magistrate
10. ponder
11. validity
12. jasmine
13. lasso
14. rampart
15. salutation
16. hassle
17. zealous
18. bolster
19. calliope
20. decathlon

Directions: Each row on this page has four letters. I will say a word. You look at the letters in the row and put a circle around the letter that you hear at the beginning of the word. Lets try the practice ones. (Examiner say each word twice.) Each child completes items 1 - 20.

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

BEST COPY AVAILABLE

A.	n	t	c	x
B.	n	r	m	g
1.	n	s	t	m
2.	t	f	c	n
3.	t	n	l	w
4.	s	r	y	t
5.	l	v	g	t
6.	r	k	n	s
7.	h	n	c	m
8.	r	c	n	t'
9.	g	m	s	t
10.	r	p	d	r
11.	l	v	t	d
12.	j	s	m	n
13.	l	h	s	m
14.	m	p	t	r
15.	l	t	s	n
16.	s	l	h	k
17.	l	z	r	s
18.	l	s	t	b
19.	l	p	n	c
20.	d	c	h	t

TEST 1 -- WHOLE WORD PRODUCTION
SPECIAL RULES

BEST COPY AVAILABLE

A. hurt

B. boot

1. hopping
2. honest
3. isle
4. hawk
5. flour
6. thorn
7. churn
8. phrase
9. thing
10. either
11. patient
12. bomb
13. hasten
14. waffle
15. took
16. worth
17. door
18. bulk
19. system
20. ghetto
21. answer
22. fudge
23. due
24. exact
25. swirl
26. roar
27. salt
28. gem
29. vase
30. wrap

Directions: I would like you to read as many of these words as you can. Lets try the practice ones. (Examiner pronounce examples correctly if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

P

TEST 1 -- WHOLE WORD PRODUCTION
SPECIAL RULES

A. hurt

B. boot

1. hopping
2. honest
3. isle
4. hawk
5. flour
6. thorn
7. churn
8. phrase
9. thing
10. either
11. patient
12. bomb
13. hasten
14. waffle
15. took
16. worth
17. door
18. bulk
19. system
20. ghetto
21. answer
22. fudge
23. due
24. exact
25. swirl
26. roa[illegible]
27. salt
28. gem
29. vase
30. wrap

TEST 2 -- NONSENSE WORD PRODUCTION

A. dem

B. ont

1. bov
2. brike
3. nirp
4. strim
5. sweem
6. thoar
7. rax
8. pape
9. flaf
10. tright
11. gube
12. crepun
13. omple
14. arp
15. mund
16. erbee
17. dalt
18. goot
19. distle
20. mentat

Directions: This page has all nonsense words. That means that none of the words make any sense. They are just letter sounds put together. Lets try the practice ones. (Examiner correctly pronounce the practice words if the child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

F

TEST 2 -- NONSENSE WORD PRODUCTION

A. dem

B. ont

1. bov
2. brike
3. nirp
4. strim
5. sweem
6. thoar
7. rax
8. pape
9. flaf
10. tright
11. gube
12. crepun
13. omple
14. arp
15. mund
16. erbee
17. dalt
18. goot
19. distle
20. mentat

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS

A. five

B. boat

1. wait
2. tease
3. hire
4. dome
5. fruit
6. cove
7. pint
8. grape
9. cheek
10. sight
11. clay
12. grow
13. cute
14. leave
15. rude
16. fair
17. slice
18. bold
19. thief
20. flute

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce the practice words correctly if the child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

F

TEST 3 -- PRODUCTION OF WORDS WITH LONG VOWEL SOUNDS

A. five

B. boat

1. wait
2. tease
3. hire
4. dome
5. fruit
6. cove
7. pint
8. grape
9. cheek
10. sight
11. clay
12. grow
13. cute
14. leave
15. rude
16. fair
17. slice
18. bold
19. thief
20. flute

F

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. dog

B. fun

1.	bag	11.	rust
2.	peg	12.	want
3.	hid	13.	dip
4.	rob	14.	mast
5.	some	15.	hut
6.	scan	16.	cap
7.	sent	17.	dead
8.	drill	18.	frog
9.	flex	19.	wink
10.	lot	20.	luck

Directions: Read as many of these words as you can. Lets try the practice ones. (Examiner pronounce words correctly if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

F

TEST 4 -- PRODUCTION OF WORDS WITH SHORT VOWEL SOUNDS

A. dog

B. fun

1. bag
2. peg
3. hid
4. rob
5. some
6. scan
7. sent
8. drill
9. flex
10. lot
11. rust
12. want
13. dip
14. mast
15. hut
16. cap
17. dead
18. frog
19. wink
20. luck

F

TEST 5 -- PRODUCTION OF CONSONANT VOWEL COMBINATIONS

BEST COPY AVAILABLE

A. ak

B. fi

1. ra
2. fe
3. pa
4. we
5. ab
6. ma
7. ja
8. ko
9. ki
10. et
11. je
12. it
13. ji
14. la
15. at
16. re
17. to
18. od
19. as
20. ru

Directions: This page has two-letter items. Some make sense and some are nonsense words. Tell me what the letters say together. Lets try the practice ones. (The child may use either the long or short vowel sound. If the child cannot pronounce the examples using either sound, the examiner should pronounce them, using a short vowel sound; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

F

TEST 5 -- PRODUCTION OF CONSONANT VOWEL COMBINATIONS

A. ak

B. fi

1. ra
2. fe
3. pa
4. we
5. ab
6. ma
7. ja
8. ko
9. ki
10. et
11. je
12. it
13. ji
14. la
15. at
16. re
17. to
18. od
19. as
20. tu

TEST 6 -- LETTER SOUND PRODUCTION

A. k

B. m

b	l	h	w	p	e	a	d	v	j
y	r	t	f	n	z	u	s	o	g

Directions: I want you to tell me the sound each one of these letters make. Lets try the practice ones. (Examiner pronounce the examples if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

****

TEST 7 -- LETTER NAMING

A. j

B. p

k	v	n	i	t	q	g	u	h	b
d	s	z	y	o	r	c	w	x	m

Directions: Tell me the names of each of these letters. Lets try the practice ones. (Examiner pronounce the examples if child does not; then have child repeat the examples.) Discontinue the subtest after 6 consecutive errors.

F

TEST 6 -- LETTER SOUND PRODUCTION

A. k

B. m

b	l	h	w	p	e	a	d	v	j
y	r	t	f	n	z	u	s	o	g

****

TEST 7 -- LETTER NAMING

A. j

B. p

k	v	n	i	t	q	g	u	h	b
d	s	z	y	o	r	c	w	x	m

F

TEST 8 -- RECOGNITION OF THE VISUAL FORM OF NONSENSE WORDS

BEST COPY AVAILABLE

A. dem

B. ont

1. brike	11. cre:yn
2. strim	12. arp (harp)
3. thoar (oar)	13. erbee (germ)
4. pāpe	14. goot (soon)
5. tright	15. mentat
6. flaf	16. gūbe
7. rax	17. omple
8. swēem	18. mund
9. nirp (dirt)	19. dalt (salt)
10. bov	20. distle

Directions: This page has rows of all nonsense words. Each row has four words. I will say a nonsense word and your job is to look across the row and find the same word I said. Put a circle around the word. Lets try the practice ones. Each child completes items 1 - 20.

F

TEST 8 -- RECOGNITION OF THE VISUAL FORM OF NONSENSE WORDS

A.	deu	lem	dem	emd
B.	ent	ont	tno	onf
1.	ikber	brike	brolo	huoke
2.	stuom	rimst	strim	ulnim
3.	thuno	sluar	oarth	thoar
4.	pago	papo	apep	jope
5.	ightrt	tright	luepht	tropre
6.	fafl	hoaf	flaf	flol
7.	rax	xar	rah	nax
8.	wemes	swion	nurem	sweem
9.	pirn	ning	uerp	nirp
10.	bov	bot	vob	hov
11.	crepun	crigom	puncre	onigun
12.	arg	arp	orp	par
13.	beere	erbee	onhee	erloo
14.	toog	gonl	puot	goot
15.	mewlal	mentat	entamt	wiulat
16.	buge	gulo	gube	pobe
17.	pleom	omple	ungle	omgho
18.	mund	dumn	muoh	wend
19.	lolt	tald	dahl	dalt
20.	dinfho	lonlle	distle	isteld

TEST 9 -- RECOGNITION OF THE VISUAL FORM
OF CONSONANT VOWEL COMBINATIONS

A. ak

B. fi

BEST COPY AVAILABLE

1. fe
2. we
3. ma
4. ko
5. et
6. ki
7. ja
8. ab
9. pa
10. ra
11. it
12. la
13. re
14. od
15. tu
16. as
17. to
18. je
19. at
20. ji

Directions: This page has rows of consonant vowel combinations. I will say the sound that two letters make together and you put a circle around the two letters that make that sound. Lets try the practice ones. (Examiner pronounce all combinations with the short vowel sound.) Each child completes all items 1 - 20.

F

TEST 9 -- RECOGNITION OF THE VISUAL FORM OF CONSONANT VOWEL COMBINATIONS

A.	ab	uk	ak	mi
B.	hi	fi	fo	la
1.	fa	fe	ab	te
2.	we	ku	wo	ne
3.	tu	ba	ma	mo
4.	mo	ko	ka	ut
5.	it	vo	ev	et
6.	su	ko	wi	ki
7.	ju	ja	em	pa
8.	ja	ob	ab	ad
9.	pa	pe	ba	lu
10.	ra	sa	ok	ru
11.	wu	it	im	ut
12.	lo	fo	la	ra
13.	lu	re	fe	ru
14.	ob	od	id	ta
15.	tu	fi	to	su
16.	am	it	is	as
17.	mo	ru	to	ta
18.	wi	ja	je	fe
19.	as	at	ok	it
20.	ji	ju	ma	fi

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

A. modification

B. versatility

1. hilarious
2. seculsion
3. ruffian
4. dainty
5. zeppelin
6. bastion
7. camphor
8. lacquer
9. jetsen
10. jocular
11. kilometer
12. vociferous
13. palpitate
14. goiter
15. yield
16. mezzanine
17. winch
18. facade
19. tolerant
20. nocturnal

Directions: Each row on this page has four letters. I will say a word. You look at the letters in the row and put a circle around the letter that you hear at the beginning of the word. Lets try the practice ones. (Examiner say each word twice.) Each child completes items 1 - 20.

F

BEST COPY AVAILABLE

TEST 10 -- RECOGNITION OF INITIAL LETTER SOUNDS

A.	f	d	m	t
B.	s	t	l	v
1.	h	l	r	s
2.	p	s	c	l
3.	r	f	n	p
4.	n	d	m	t
5.	l	n	z	p
6.	n	b	s	t
7.	p	h	r	c
8.	c	l	r	t
9.	j	n	s	t
10.	c	j	l	r
11.	l	m	k	t
12.	s	r	v	c
13.	p	l	t	b
14.	f	r	t	g
15.	l	y	d	w
16.	s	n	m	z
17.	n	w	h	c
18.	f	c	d	v
19.	r	l	n	t
20.	c	n	t	r

F

Learning Hierarchies in Reading Comprehension

Abstract

Learning hierarchies in reading comprehension have not been widely investigated. An array of tasks was constructed to measure skills in reading comprehension. Tasks were administered individually to 24 good and 24 poor readers at grade 2 reading level. Differential difficulty among the tasks were assessed by analysis of variance and Guttman scaling analysis. Results were that a learning hierarchy (Guttman scale) for good readers consisted of: decoding, sentence listening comprehension, sentence reading comprehension, and sentence question answering. Poor readers had a different hierarchy (Guttman scale) of: sentence reading comprehension, decoding, sight vocabulary and sentence question answering. Reproducibilities were .94 and 1.00 respectively. Implications for reading acquisition in good and poor readers and instructional development are discussed.

Learning Hierarchies in Reading Comprehension

Investigation of learning hierarchies occupies a prominent position in the psychology of instruction. According to Gagné (1973) who has been the principal architect of this conceptual framework, a learning hierarchy is defined as a "description of successively available intellectual skills. Each skill. . . is placed in a hierarchy in such a way that the skills subordinate to a given skill. . . contribute substantively to the learning of the given skill, in the sense of exhibiting positive transfer to it."

Two criteria have been used to determine whether intellectual skills are hierarchically related. First, the difficulty of the skills is assessed. If two skills, A and B, are logically associated and if A is consistently more difficult than B, the two skills may form a hierarchy. A popular procedure for analyzing the consistency of the order of difficulty between (among) skills is Guttman scaling analysis. In the words of Glaser and Resnick (1972), "scaling data indicate the extent to which performance on lower order tasks can reliably be predicted from information concerning performance on higher order tasks." When a set of tasks is scalable, subjects who can peform a higher order task can perform all lower order tasks; and subjects who fail a given lower order task fail all higher order tasks.

The second criterion is positive transfer among tasks. If learning task B facilitates the acquisition of the more difficult task A, a hierarchy is said to obtain. Learning hierarchies in mathematical skills have been shown by Gagné (1965) to fulfill the second criterion as well as the first. It should be noted that scaling data have only suggestive utility for instruction purposes, since these data do not necessarily imply positive transfer in the array of tasks or skills. However, scaling studies have heuristic value for the identification of tasks which may be tested in transfer experiments. Successful demonstration of positive transfer provides a hierarchy which may yield a sequence of objectives that can be used in the design and evaluation of curricula (Glaser, 1973).

In reading, the need for well documented learning hierarchies is urgent. There is a plethora of commercially available instructional approaches that are based on hypothetical hierarchies or no hierarchies at all (Aukerman, 1971). Recent studies (Hardy, et. al 1972; Oliver et. al. 1972); Samuels, 1969; and Samuels, 1973) have illustrated the importance of hierarchies in the acquisition of grapheme-phoneme correspondences. At a more molar level Guthrie (1973) suggested that a wide range of grapheme-phoneme association skills (decoding) may be hierarchically related. An array of skills embracing single letter sound production and fourth grade level word reading were found to be consistently ordered in difficulty.

Transfer tests of skills in that array are needed to determine whether they meet the second criterion for hierarchies.

The issue to which the present study addresses itself is whether a set of skills that are thought to be related to reading comprehension are arranged in a learning hierarchy. Although both criteria presented previously are relevant to the determination of whether skills are hierarchically related, this paper examines the first criterion, i.e. differential difficulty. Skills of interest for this issue include: oral reading (decoding), sentence listening comprehension, sight vocabulary, sentence reading comprehension and answering questions on sentences. Do these skills meet the criteria for learning hierarchies or not? Cases have been made for both sides of the argument. For example, Gough (1972) contends that reading must consist of: the visual perception of written characters, mapping the characters onto speech-like strings of phonemes, and retrieval of meanings with phonemic cues. He believes that these processes are necessary since a parsimonious cognitive system would not allow the formation of the complex rules that are needed to relate English orthography directly to word meanings. On the other side, many authors suggest that "visual reading" may be possible in which meaning is derived directly from print without the intermediate stage of decoding to speech (Baron, 1973; Cohen, 1968; Huey, 1908; Kolers, 1970). In terms of the present study, this debate relates to the issue of whether decoding (oral reading) is necessary for and hierarchically related to sentence comprehension.

A second question raised in this study is whether a learning hierarchy that may be identified for normal readers is also observed in poor readers. If a hierarchy is observed in normal readers, do poor readers manifest the same hierarchy, a different one, or none at all?

Method

Subjects

Two groups of subjects were used in this study. A normal group of 24 second graders were included whose chronological age was 7 years 1month. The mean Peabody Picture Vocabulary Test IQ of this group was 107.90 and the mean Gates-MacGinitie Comprehension Level A reading grade level was 2.30. The poor readers were selected from a non-graded school for special reading instruction. This group of 24 Ss had a mean Peabody Picture Vocabulary Test IQ of 103.04 which did not differ significantly from the normal group ($t = 1.74$, $df = 46$, $p > .05$). Their mean Gates-MacGinitie Comprehension Level A, B and C performance was 2.27 which did not differ significantly from the normals ($t = .12$, $df = 46$, $p > .05$). The chronological age of this group was 10 years 4 months, significantly higher then the comparison group ($t = 15.49$, $df = 46$, $p < .01$).

To insure that the poor readers did not have other outstanding deficiencies additional tests were given. The WISC Full Scale of the group was 102.50, with a verbal IQ of 99.13 and a performance IQ of 105.92. On the Wepman Test of Auditory Discrimination the mean number of errors was 3.27 out of 30 items which is judged to be "adequate" based on the norms. The Templin-Darley Articulation Test revealed that 22 children were normal and 2 had mild difficulties not requiring speech therapy. The auditory memory was 7-2 mental age equivalent based on the digit series test from the Binet. Visual memory as measured by the visual sequential memory subtest of the Illinois Test of Psycholinguistic Abilities showed a mental age equivalent of 7-11. These latter two abilities are both lower than the chronological ages of the Ss indicating that general memory skills were low in this group.

Materials

Two stories were selected from Basal readers (New Basic Readers by Scott-Foresman; Bank Street Reader). The stories were shortened to 400 words and were at the 3^1 difficulty level according to the Spache (1953) formula.

The observations in the study will relate to whether there is a hierarchial relationship among the tasks. Each task requires a combination of knowledge (previously acquired information) and skill (sequence of cognitive operations), in verying proportions. For this report the term skill is used to refer to all cognitive functions that are needed to perform a given task. If tasks are ordered in difficulty, it will be inferred that cognitive skills needed to perform the tasks are ordered in strength or availability. Some of the skills are complex and may include many subskills. However, the integrity of any one skill will be supported if it provides interpretable and replicable data.

Tasks

Sight vocabulary (SV). A test of 40 items was constructed by selecting single words that were not functors (prepositions, articles, conjunctions) from one of the stories (story 1). Each item had 4 alternatives: a correct answer that was synonomous with the stem, an incorrect item that was visually similar to the stem, and two other incorrect alternatives. A large majority (93 percent) of the alternatives were on the Dale-Chall list of 769 easy words, making them lower than third grade level. An example item was: street (as the stem); wall, road, house, string (as the alternatives). Clearly, the item could be successfully answered with a simple comparison of word meanings between the stem and the alternatives. Fine semantic discriminations were not required. The test contained this item and Ss worked silently, attempting to circle correct answers without assistance.

Sentence reading comprehension (SRC). Story 1 was used to construct this task. The maze task, previously used by the first author (Guthrie, 1973b) was employed. In the story 40 maze items were created by replacing about every tenth word with 3 alternatives. The alternatives included the original word as the correct answer, an incorrect alternative that was semantically anomalous, and an incorrect alternative that was syntactically anomalous. The latter alternatives were drawn from other words in the story to maintain a consistent level of difficulty. Here is an example.

stand
People in the street would nine and stare as Baby
wear

floor
Hattie's carriage passed because they thought at first that
see

if
Hattie's mother had a radio hot the carriage.
in

Four categories of form classes were used. An equal number (10) of items were formed around words that served as nouns, verbs, modifiers, and function words in the passage. The class of nouns included nouns and pronouns; verbs included transitive, intransitive and auxiliary words; modifiers included adjectives and adverbs; functions included prepositions, articles and conjunctions. The Ss completed a short example story with the examiner and then worked silently attempting to circle the correct alternatives in the passages, which were presented in typed form on 8 1/2" by 11" paper.

<u>Sentence question answering (SQA)</u>. This task was constructed from story 1 and contained 20 questions which were

divided into 4 sets of 5 questions. Each set was placed after a section consisting of about one-fourth of the passage. Questions following a section were relevant to that section, thus, Ss read a 100 word section, answered questions over it, read the next 100 word section, answered questions over it and repeated the cycle until the questions were completed. It is evident that memory is one cognitive function that is demanded in addition to comprehension in this task. Since each question was based on only one sentence and did not require inter sentence processing the task is termed sentence question answering.

The questions were divided into four types using the Bormuth, Carr, Manning and Pearson (1970) classification system that includes: rote, semantic substitution, syntactic transformation and compound questions. An equal number of types were distributed throughout the group of 20 items. Each question was based on one sentence or a clause within a sentence in the passage. The rote questions were created by converting a sentence to a question with no change in the sentence except the insertion of an interrogative word. For example, one sentence was: "The ear muffs almost stopped the sound of Hattie's squeak." The question was: "The ear muffs almost stopped what?" The alternatives for the question were: mother and father, the cotton, the sound, Hattie. Children were requested to circle the most appropriate answer. Alternatives were selected from words in the sentence or adjacent sentences.

The semantic substitution questions were constructed by replacing the original words or phrases in the sentence with words or phrases that have highly similar meanings and inserting an interrogative word. Alternatives were semantic substitutions for words or phrases in the sentences. For instance, one original sentence was: "After a while Hattie's talking began to make trouble for her." The question was: The girl's speaking started what?" The four alternatives were: sleep, a party, problems, talk.

The syntactic transformation items were developed by generating a syntactic transformation of the sentence that did not change its basic meaning and inserting an interrogative word. The original words of the sentence were retained and the alternatives were drawn from the sentence or adjacent sentences. An example of an original sentence is: "At first, Hattie's father and mother were very proud of their daughter." The question was: "The daughter made whom very proud?" The alternatives consisted of: mother and father, grown ups, mother, people. The compound items consisted of the operations that were performed for the syntactic transformation and semantic substitution items. An example of a sentence from the passage was: "Hattie talked so long that people who stopped to listen to her would be late for school or late for work or late for supper." The question was: "Men and women would not be on time because of what?" The alternatives were:

jobs and food, a girl speaking, their children, cars and trucks.

Oral reading (OR). This task was constructed from story 2. Words from the sentence reading comprehension (SRC) task in story 2 were used. All words in the task, including the alternatives, were arranged in a haphazard order in a list with no duplications of words. The Ss read the list, which contained 196 words, aloud with no assistance. The E judged each word in terms of whether it was unmistakably similar to the same word in the Ss oral vocabulary.

Sentence reading comprehension (SRC). This task was constructed for story 2 as well as story 1 to facilitate comparison among all of the tasks. The maze procedure was used in a manner identical to that used for story 1, and Ss performed the task similarly. There were 40 items which included equal numbers of nouns, verbs, modifiers and functions with 3 alternatives for each maze item.

Sentence listening comprehension (SLC). This task employed the same materials as the SRC for story 2. The story and maze items were identical. The presentation, however, was aural rather than visual. The Ss had answer sheets that had 40 3-choice written alternatives. They responded by attempting to circle the correct alternative. This response mode was identical to that used for SRC of story 2. The story was read aloud to the Ss by E with the signal of a pencil tap occurring before and after each maze item. That is, E read the words in a sentence prior to a maze item, tapped the pencil, read the maze alternatives, tapped the pencil, read the words following that item to the point immediately preceding the next item. The Ss then responded by circling an answer. The E repeated each of the presentations up to 3 times if the Ss requested it to minimize demand on memory. The entire story was presented in this manner.

Procedure

Preliminary information on reading level was gathered by administering the Gates-MacGinitie comprehension test to both good and poor readers. The Peabody Picture Vocabulary Test (PPVT) was given to the good readers individually and recent scores on the test were gathered from the cumulative records of the poor readers. The oral reading task was given individually to all Ss. Other tasks were given in groups of 3-4 in a small, quiet room. The sight vocabulary, sentence question answering and sentence reading comprehension (story 1) were administered in counter balanced order such that all possible orders of tasks were given an equal number of times. The oral reading, sentence listening comprehension and sentence reading comprehension (story 2) tasks were administered in counter-balanced order following the first set. For the normal readers all 6 tasks were administered in 3 sessions lasting approximately 50 minutes each with a rest at the midpoint of each session. For the poor readers, 4-5 sessions of 50 minutes were needed.

Unlimited time was provided for the administration of all tasks and Ss were encouraged to complete every task.

Results

Before the primary issues of the study will be addressed, the reliabilities and intercorrelations of the measures will be presented. For all tasks internal consistency reliabilities were computed with the KR21 formula. Outcomes of these calculations are in Table 1. These data indicate that the abilities for all measures were above .80 except for the sentence question answering reliabilities which were .73 for poor readers and .79 for good readers, adequate for purposes of this study.

Intercorrelations among all the tasks are presented in Table 2. These correlations were uniformly high for normal readers. For poor readers the correlations were substantial except that sentence listening comprehension did not correlate significantly with any other variable except chronological age. Note that the r for chronological age and sentence listening comprehension was probably not significant for normals due to noticeably restricted range in this group. Validity of the sentence question answering task may be determined from its correlation with the Gates-MacGinitie Comprehension test. For normals it was .78 and for poor readers it was .41, both of which are significantly different from zero. Thus the most complex of the tasks is clearly associated with standardized comprehension test performance.

Recall that two sentence reading comprehension tasks were administered based on different basal stories. Performance on these two measures was found to be similar ($t = .013$, $df = 46$, $p > .50$) indicating that the difficulty of the 2 passages was equal. The equivalence in sentence reading comprehension permits comparison across the 5 tasks which were based on different passages, in some instances, as outlined in the method section. Subsequent analyses were based on the mean of the two sentence reading comprehension tasks for all Ss.

The primary focus of the study was on the differential difficulty of the tasks for the 2 groups of subjects. Since some tasks had three multiple choice alternatives and others had four, the raw scores were corrected for guessing with the formula $R = R - \frac{W}{n-1}$ (Cronbach, 1970). The correction was not made for oral reading since the probability of correct responses from guessing cannot be accurately estimated. An analysis of variance was conducted using a 2 (groups) x 5 (tasks) design with repeated measures on the second factor. The results were that the poor readers obtained higher scores on all of the tasks combined than normal children with whom they were matched on reading comprehension ($F = 13.88$, $df = 1/46$, $p < .01$). There were significant differences among the tasks ($F = 40.50$, $df = 4/184$, $p < .01$) and there was a significant group x task interaction ($F = 8.18$, $df = 4/184$, $p < .01$) Post-hoc tests on these data using the Neuman-Kuels procedure are reported in Table 3. It is evident that the rank order of the tasks was different for normal and poor readers. For normals, oral reading was superior to all other tasks with a mean of 61.08 percent correct. Sentence listening comprehension and sentence reading comprehension were not significantly different with a combined mean of 37.90 percent correct. Finally, lowest on the hierarchy were sentence reading comprehension, sentence question answering and sight vocabulary which had a combined mean of 27.36 percent correct. For poor readers, however, a different picture emerged. These skills, sentence listening comprehension, oral reading and sentence reading comprehension were similar in difficulty and were superior to other skills with a combined mean of 71.36 percent correct. Sight vocabulary was significantly lower than the previous group at 54.75 percent correct. Finally, sentence question answering was significantly lower than the others with a level of 36.08 percent correct.

In learning hierarchy studies, Guttman scaling analysis (Nunnally, 1967) is the customary procedure for determining the extent to which performance on lower order tasks may be predicted from performance on higher order tasks. Guttman analyses were conducted on the data for normal and poor readers using the following procedure. For each group a grand mean for all individuals on all tasks was calculated. The performance of each individual on each task was then dichotomously scored depending on whether the score was above or below the grand mean for the particular group. These data were then submitted to Multiple

Scalogram Analysis, a Guttman analytic procedure that extracts more than one scale from the data if it is present (Lingoes, 1963).

Outcomes of this analysis are presented in Table 4. For normal readers, the scaled tasks were ranked from relatively easy to relatively difficult as follows: oral reading, sentence listening comprehension, sentence reading comprehension and sentence question answering. This is the same rank order as was observed for the group means and demonstrates that the hierarchy which was observed for the group also held, with a high degree of probability, for all individuals within the group. The reproducibility of this scale was .94. For poor readers, the items in the scale also paralleled the group means, ranging from relatively easy to relatively difficult as follows: sentence reading comprehension, oral reading, sight vocabulary and sentence question answering. The reproducibility for this scale was 1.00, showing that all tasks conformed to the scale for all Ss. Note that some tasks scaled even though they were not significantly different on the group data. This outcome represents unique information provided by the Guttman analysis. It indicates that scores of individuals on the tasks fell into the pattern defined by the Guttman scale although the distribution of some tasks overlapped considerably.

Discussion

Presence of learning hierarchies in reading comprehension is supported by this investigation. The array of skills including oral reading, sentence listening comprehension, sentence reading comprehension, sight vocabulary and sentence question answering. were found to have differential levels of strength. The 5 tasks were significantly different in difficulty based on data for normal and poor readers. Furthermore, 4 of the tasks were arranged in a Guttman scale for the two groups separately. For the normal readers the tasks which entered the Guttman scale ranged from easy to difficult as follows: oral reading, sentence listening comprehension, sentence reading comprehension and sentence question answering. The reproducibility was .94, indicating that if an individual was relatively proficient in sentence question answering, he was likely to be proficient in all of the lower order skills. Conversely, if an individual was relatively low on a given skill, e.g. sentence listening comprehension, he was relatively low on all higher order skills, with a high degree of probability. Poor readers manifested a qualitatively different learning hierarchy than normal readers. Skills that entered the Guttman scale for poor readers ranged from easy to difficult as follows: sentence reading comprehension, oral reading, sight vocabulary, sentence question answering. The reproducibility was 1.00 showing that the Guttman scale obtained for all Ss on all tasks.

Differences between normal and poor readers included: 1) components of the learning hierarchies, 2) rank order of the components, and 3) relative degree of proficiency among components. The scale contained sentence listening comprehension for normals but not for poor readers suggesting that listening skills are not systematically related to reading skills in poor readers. The low intercorrelation between sentence listening comprehension and other reading skills supports this notion. Perhaps this indicates that comprehension difficulties of poor readers stem from their failure to generalize language skills from listening to reading. However, it may mean that listening and reading are not as inherently related as popularly believed for disabled readers. The dependence of reading comprehension on listening comprehension for normal readers is substantiated by this study.

Sight vocabulary was present in the learning hierarchy for poor readers and not for normals. Recognition of printed word meanings is systematically related to reading comprehension, particularly answering questions on sentences in poor readers. Among normals, sight vocabulary and answering questions on sentences were similar in difficulty and were moderately correlated. Under these correlations, absence of sight vocabulary from the scale, indicates that for individuals who had a score on one test (e.g. sight vocabulary) above the mean and a score on another test (e.g. sentence question answering) below the mean, the distribution of scores around the mean was random.

A marked distinction between the groups derives from the fact that oral reading in normals was substantially more proficient than the next skills in the hierarchy, sentence reading comprehension and sentence listening comprehension. In contrast, the oral reading level of poor readers was the same as their sentence reading comprehension level. A possible interpretation for this is that poor readers do not learn decoding skills easily. Their sentence reading comprehension skill develops to the point where they can comprehend most of what they can read orally. One impediment to their improvement in sentence comprehension may be lack of decoding fluency.

Normal readers answered questions on sentences at the same degree of proficiency as they comprehended sentences suggesting generalization from lower order to higher order skill. However, poor readers were substantially lower (about 50 percent) in sentence question answering than sentence reading comprehension. These results suggest that the additional cognitive demands of memory, processing questions and relating questions to sentences pose greater cognitive challenges for poor readers than normal readers.

It was observed in Table 3 that poor readers had a higher grand mean than normal readers. Recall that the groups were initially matched on the Gates-MacGinitie Comprehension test. It is comforting that they were also similar (not significantly different) on the sentence question answering task. However, note that sentence question answering is significantly lower than other skills on the hierarchy for poor readers. It is possible that matching poor readers with normal readers on this task, produced an arrangement in which skills that are prerequisite to this task were operating at higher levels of proficiency for poor than good readers.

A question may be raised regarding the effect of instructional method in learning hierarchies. It is possible that one teaching approach, e.g. phonics oriented, may produce one learning hierarchy and a second teaching approach, e.g. sentence comprehension oriented, may produce a second learning hierarchy. In the present study the normal children received an eclectic, comprehensive reading program including word recognition, phonics, sentence reading and listening comprehension activities. The poor readers had also received diverse instruction as a group. For some children phonics had been emphasized and for others sentence comprehension had been emphasized. Since the reproducibility of the hierarchy was so high (1.00) for the poor readers, there is no evidence that the different teaching emphases produced different hierarchies. Although it appears that the hierarchies will not be drastically modified by different teaching programs, the issue remains open to empirical investigation.

Learning hierarchies are confirmed when the two criteria of differential difficulty and positive transfer are met. This study suggests that differential difficulty is present for several skills related to reading comprehension, and that experimental tests of positive transfer among the components in the hierarchies is warranted. The transfer experiments are needed before implications for curriculum development may be drawn confidently. However, if learning hierarchies in comprehension that meet both criteria are qualitatively different for good and poor readers, an important empirical issue for curriculum development arises. Should curricula be tailored and matched to the hierarchies of different types of children or should curricula be organized to parallel the hierarchies of efficient learners, with the implication that hierarchies of poor readers should be modified rather than accomodated?

The implication of a learning hierarchy for instruction is that teaching will be optimized by progressing from lower order to higher order skills. In the present study, a hierarchy of skills was found poor readers with materials at the 3^1 level. One plausible instructional approach derived from this hierarchy may be termed a hierarchical mastery procedure. For any given level of materials, e.g. 3^1, the lowest order skill should be mastered first. Then the next skill in the hierarchy should be mastered. When the highest order skill has been acquired, the next level of difficulty, e.g. 3^2, may be introduced. The lowest order skill may be taught to mastery at this new level of difficulty and the cycle may be repeated until all skills at the 3^2 level are attained. Such a prodedure may prove more efficient than instruction in which learning hierarchies are ignored and skills are taught in a nonsystematic order.

Footnotes

1. Project support was generously provided by the Spencer Foundation and The Joseph P. Kennedy, Jr. Foundation.

2. Assistance was furnished by the principal and staff of the Kennedy School and Our Lady of Victory School, Baltimore, Maryland. Their cooperation is gratefully recognized.

References

Aukerman, R. Approaches to Beginning Reading. New York, John Wiley & Sons, Inc., 1971.

Baron, J. Phonemic stage not necessary for reading. Quarterly Journal of Experimental Psychology, 1973, 25, 241-246.

Bormuth, J.R., Carr, J., Manning, J., Pearson, D. Children's comprehension of between and within sentence syntactic structures. Journal of Educational Psychology, 1970, 61 (5), 349-357.

Cohen, G. A comparison of semantic,acoustic and visual criteria for matching of word pairs. Perception & Psychophysics, 1968, 4 (4), 203-204.

Cronbach, L. Essentials of Psychological Testing. New York, Harper & Row, 1970.

Gagne, R.M. & Staff, Some factors in learning non-metric geometry. Monographs of the Society for Research in Child Development, 1965, 30 (1), 42-50.

Gagne, R.M. Learning and instructional sequence.in Kerlinger, F. Review of Research in Education. Istasia, Illinois, F.E. Peacock Publishers, Inc., 1973.

Glaser, R. & Resnick, L.B. Instructional psychology in Mussen, P. & Rosenzweig, M. (Eds.) Annual Review of Psychology, 1972, 23, 207-277.

Glaser, R. Educational psychology and education, American Psychologist, 1973, 28 (7), 557-566.

Gough, P. One second of reading in Kavanagh, J. & Mattingly, I. Language by Ear and by Eye. Cambridge, Massachusetts, The MIT press, 1972.

Guthrie, J.T. Models of reading and reading disability. Journal of Educational Psychology, 1973a, 65 (1), 9-18.

Guthrie, J.T. Reading comprehension and syntactic responses in good and poor readers. Journal of Educational Psychology, 1973b, 65 (3), 294-299.

Hardy, M., Smythe, P.C., Stennett, R.G. & Wilson, A.R. Developmental patterns in elemental reading skills: phoneme-grapheme and grapheme-phoneme correspondence. Journal of Educational Psychology, 1972, 62 (5), 433-436.

Huey, E.B. The Psychology and Pedagogy of Reading. New York, MacMillan, 1908.

Kolers, P. Three stages of reading in Levin, H. & Williams, J. Basic Studies on Reading. New York, Basic Books, Inc., 1970.

Lingoes, J. Multiple scalogram analysis: a set theoretic model for analyzing dichotomous items. Educational and Psychological Measurement, 1963, 23 (3), 501-524.

Nunnally, J.C. Psychometric Theory. McGraw-Hill, New York, 1967.

Oliver, P.R., Nelson, J.M. & Downing, J. Differentiation of grapheme-phoneme units as a function of orthography. Journal of Educational Psychology, 1972, 63 (5), 487-492.

Samuels, S.J. Effect of simultaneous versus successive discrimination training on paired associate learning. Journal of Educational Psychology, 1969, 60 (1), 46-48.

Samuels, S. Jay Effect of distinctive feature training on paired-associate learning. Journal of Educational Psychology, 1973, 64 (2), 164-170.

Spache, G. A new readability formula for primary grade reading materials. The Elementary School Journal, 1953, 53, 410-413.

TABLE 1

Means, Standard Deviations and

Reliabilities of Tasks in Reading

Comprehension Learning Hierarchy Study

Group and Task	Mean Percent Correct[b]	Standard Deviation	Reliability[c]
Normal Readers			
Sight Vocabulary	23.33	28.15	.89
Sentence Reading Comprehension[a]	34.17	32.13	.88
Sentence Question Answering	24.58	24.69	.73
Oral Reading	61.08	24.85	.98
Sentence Listening Comprehension	41.63	27.93	.85
Poor Readers			
Sight Vocabulary	54.75	31.54	.93
Sentence Reading Comprehension[a]	66.83	28.41	.89
Sentence Question Answering	36.08	28.17	.79
Oral Reading	69.58	18.26	.97
Sentence Listening Comprehension	77.50	17.28	.83

a Mean Percent Correct of two sentence reading passages combined
b. Percent correct using raw scores which were corrected for guessing
c Reliabilities based on the KR21 formula

TABLE 2

Correlations Among the Five Tasks,

Gates-MacGinitie Comprehension and Chronological Age

for Normal and Poor Readers[(a)]

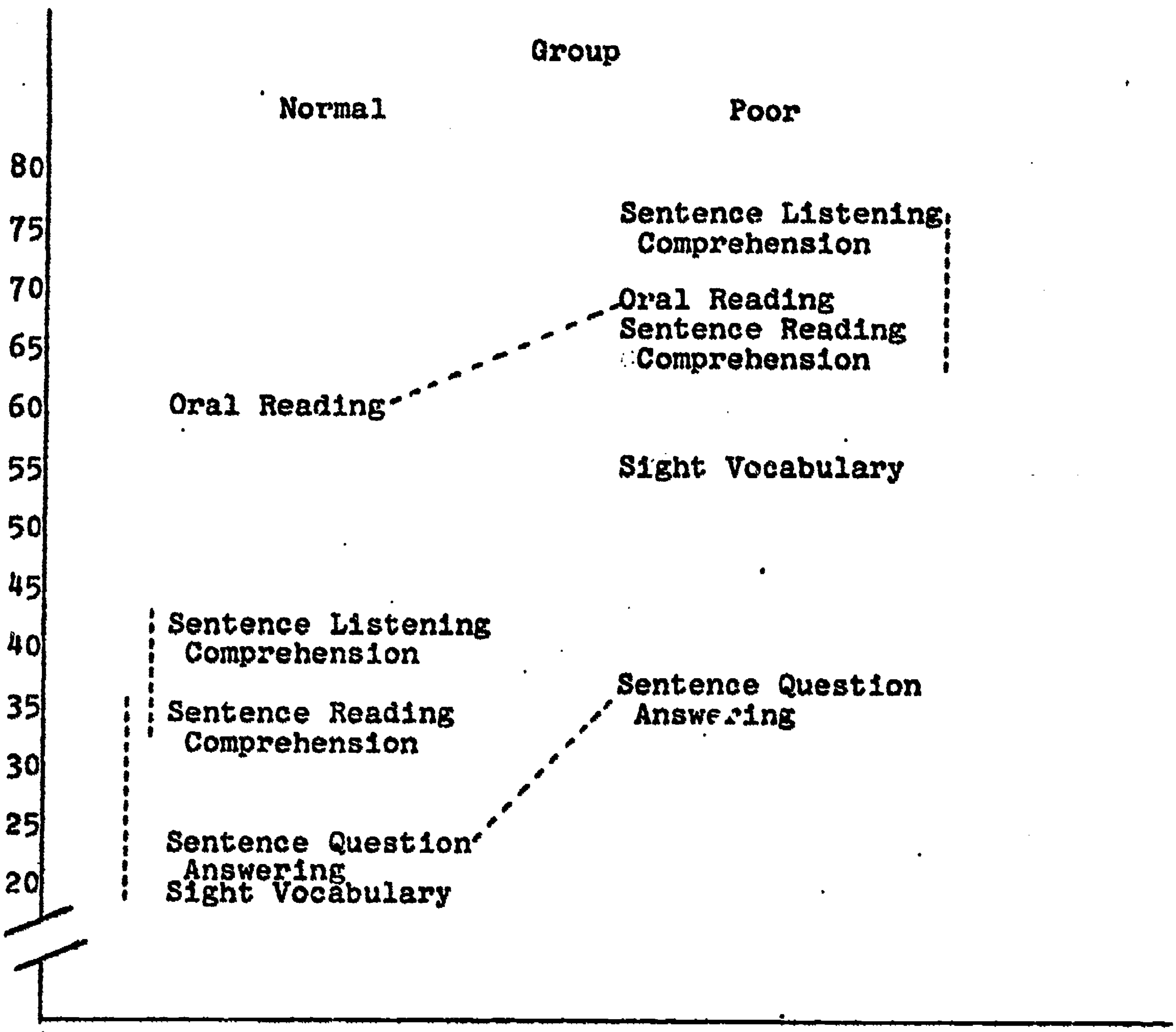
	CA	GMC	SRC	SLC	SQA	SV	OR
Chronological Age (CA)		.047	-.039	.430*	-.026	-.174	-.115
Gates-MacGinitie Comprehension (GMC)	.079		.631**	.195++	.410*+	.569**	.560
Sentence Reading Comprehension (SRC)	.193	.844**		.315++	.540**	.855**	.905
Sentence Listening Comprehension (SLC)	.034	.861**	.810**		.235+	.288	.213
Sentence Question Answering (SQA)	.024	.782**	.741**	.759**		.465*	.582
Sight Vocabulary (SV)	.154	.747**	.870**	.692**	.736**		.888
Oral Reading (OR)	-.076	.854**	.771**	.792**	.664**	.692**	

(a) Correlations for poor readers are in the upper and right half of table. Correlations for normal readers are in the lower and left half of the table.
* indicates $p < .05$
** indicates $p < .01$
\+ indicates $p < .05$ difference between poor readers and normal readers
++ indicates $p < .01$ difference between poor readers and normal readers

TABLE 3

Learning Hierarchies in Reading Comprehension

for Good and Poor Readers



Note: Entries joined by a line do not differ at $p < .01$.

TABLE 4

Guttman Analysis of

Reading Comprehension

Learning Hierarchies

Task Order

Normal Readers	Poor Readers
Oral Reading	Sentence Reading Comprehension
Sentence Listening Comprehension	Oral Reading
Sentence Reading Comprehension	Sight Vocabulary
Sentence Question Answering	Sentence Question Answering

Reproducibility

Normal Readers	Poor Readers
.94	1.00

Cornell Scores

Number of tests passed	Number of subjects: Normal	Number of subjects: Poor
4	5	5
3	4	8
2	3	2
1	6	0
0	6	9
	24	24

THE MAZE TECHNIQUE FOR ASSESSING AND MONITORING READING COMPREHENSION

John T. Guthrie

Mary Seifert

Nancy Burnham

Ronald I. Caplan

Kennedy Institute

Johns Hopkins University

THE MAZE TECHNIQUE FOR ASSESSING
AND MONITORING READING COMPREHENSION

It is widely believed that there are many different skills involved in reading and that these skills develop simultaneously in most children. In the general population, a third grade child who is relatively superior in phonic skills is likely to be relatively superior in reading comprehension. At the same time, a child who is worse than his peers in sight vocabulary is likely to be weak in understanding sentences and paragraphs. These relationships are reflected in the high positive correlation of about .70 between word identification tests such as the Wide Range Achievement Test (WRAT) and comprehension measures such as the California Reading Test (Olridge, 1964, Washington & Taska, 1970).

On the other side of the coin, there are many individual children who represent exceptions to these generalizations. Some children may be able to identify words accurately without being able to understand sentences or paragraphs. These exceptions are not inconsistent with the .70 correlation between word identification and paragraph comprehension. If the correlation coefficient between two variables is squared, the proportion of variability in one variable may be estimated. For example, .70 squared is .49 which indicates that about half of the differences between children in paragraph comprehension may be traced to word identification ability. The other half must be related to other factors such as linguistic competence, vocabulary level and memory skills.

Since it cannot be assumed that children comprehend material that they can read orally, two problems arise. The first problem is the identification of individual children whose comprehension is deficient and the identification of the comprehension level of a class of children whether they are normal, educably retarded or learning disabled.

A second problem is that teachers and reading specialists need a simple, accurate means to monitor the progress of children during the course of a reading program. Particularly if the program emphasizes comprehension skills, the comprehension levels of an individual or a group should be assessed regularly to supply feedback to the teacher about the effectiveness of the instructional approach. Standardized tests are insufficient for this purpose since they require time and money and cannot be given with sufficient frequency to provide the feedback that is needed for continuous revision and improvement of the teaching program.

The solution to these two problems lies in the use of a technique to measure comprehension levels within the classroom. There are at least three general requirements of any technique to fulfill this purpose. First, the teacher must be able to construct the measuring device using materials that are familiar to her, immediately available and useful for teaching. This requirement eliminates standardized tests as suggested previously. Second, the technique must be reliable. This is, a child should perform as well one day on the test as he does the next. The child's performance on two sections of the test which attempt to measure the same skill should be similar. Third, the technique must be valid. Use of the teacher-made technique to measure comprehension should produce similar results to use of a standardized test.

The most common form of evaluating reading comprehension consists of asking questions. After reading a paragraph or passage, the child is presented questions in written or oral form which he attempts to answer. Many difficulties attend this method of evaluating comprehension. First, if a child fails to answer one or more questions, it is difficult to determine whether he did not understand the question, did not understand the passage, failed to relate the question to the proper section of the passage or some combination of these.

A second problem associated with the use of questions is that questions vary immensely in nature and quality. Although "rote" questions are usually easier than "inference" questions, there are many types of both "rote" and "inference" questions. Some "inference" questions such as "Was it a cold morning?" that relate to the sentence "Ted pulled the wagon over the frosty grass." may be easy. This type of question may be as easy as a "rote" question that inquires "who pulled the wagon" when the story stated that "Ted pulled the wagon." Since the distinctions between "rote" and "inference" questions are complex, "rote" questions are not always easier than "inference" questions and the difficulties of different "inference" questions may vary dramatically. Suppose a teacher attempts to pose ten "inference" questions on two occasions (2 Fridays). A child's apparent comprehension on the two occasions may be higher or lower due to inconsistency in the difficulty of the questions rather than increases or decreases in his understanding of the reading material.

A third difficulty in using questions to evaluate comprehension is that the child's ability to construct answers to questions may not reflect his actual understanding of what he has read. Suppose a child is required to provide a short written answer to the question "What do we know about the day" when the passage contained the sentence "Ted pulled the wagon on the frosty grass." The child may not be able to generate the phrase "cold morning" although he would answer "yes" to the question of whether it was a cold morning. Thus the child could recognize a correct answer although he could not produce one. In this case, the response mode rather than conceptual complexity determined the difficulty of the question. Questions must be very carefull, ontrolled for type, and response mode if they are to be accurate, consistent indications of comprehension. Since such control is time consuming and impractical in many teaching situtations, the informal assessment of reading comprehension is usually less accurate and consistent than desirable.

One possible technique that may be used to measure and monitor comprehension in reading is the maze procedure. The maze procedure consists of a series of sentences which may be extracted from any story or book. The text is modified by substituting 3 alternative words for every fifth or tenth word in the story.

Here is an example:

	on		some
The truck was full of	corn .	The farmer and his truck	swam fast.
	roads		went

The child reads the material silently and circles the alternatives which he believes are correct. The number or percentage that the child circles correctly indicates the level of his comprehension of that passage. For example, suppose a child was given a 100 word passage with 20 maze items.

If he answered 15 items correctly, it could be said that he understood the passage with 75 percent proficiency.

The validity and reliability of this technique have been described in a previous article (Guthrie, 1973). The correlation between performance on 7 passages of the maze task, 160 words in length, and the Gates-MacGinitie Comprehension Subtest for both normal and poor readers was found to be .82. Clearly, the technique identifies children to be good or poor readers in a manner very similar to the well established Gates-MacGinitie Test. Relaibility of a single 160 word passage was found to be over .90. This shows that performance on short passages is likely to be internally consistent and will probably be similar across short periods of time. The experience of the author is that when the technique has been used by teachers or reading specialists to monitor the progress of children in comprehension, the accuracy of an individual child on a certain level of material, e.g. 2^2, seldom varies more than about 10 percentage points over a period of 2 - 3 days. Of course, if a child performs at 50 percent accuracy on 2^2 material one month, he may learn to comprehend at that level and improve to 90 percent accuracy 2 months later.

To examine the usefulness of the maze technique for identifying comprehension levels in normal and poor readers, an experiment was conducted in which two tasks were given to two groups of children. The subjects were 13 normal second graders who were taught with an eclectic approach including phonics, sight vocabulary, reading comprehension, and written expression. The poor readers were 13 children who attended a a special school for reading remediation. The methods for these children included elements similar to the normal group including phonics, sight vocabulary, reading comprehension and less emphasis on written expression. The primary difference between the programs was that the poor readers were taught in small groups of 4 -5 which permitted more individualization of instruction than was possible with the groups of 35 - 40 in which the normal children were placed. The poor readers were screened to insure that each child had the following characteristics: WISC IQ higher than 90, normal hearing, normal vision, normal emotional adjustment, absence of gross neurological dysfunction, and 1.5 or more years deficiency in reading. The normal children were 7.60 years old and the poor were 9.61 years old, a significant difference ($t = 6.28$, $df = 25$, $p < .005$). The Peabody Picture Vocabulary Test IQ Scores of the normal and poor groups were 104.66 and 104.15 respectively. The reading levels of the normal and poor groups were 2.73 and 2.50 respectively on the Metropolitan Achievement Test - Word Knowledge. These levels were not significantly different.

The materials consisted of a 160 word passage from stories in basal readers which were at the 2^1 grade level according to the Spache (1953) formula. In the oral reading task (task 1) the children were required to read all of the words aloud that were in the passage. The words were presented one at a time in a random list and the child's ability to pronounce each word was assessed. In the maze task, (task 2) the passage was modified according to the procedure described previously. About every fifth word was replaced with 3 alternatives which included: a) the correct word; b) an incorrect word that was the same form class (verb, noun, function word, modifier) as the correct word; and c) an incorrect word that was a different form from the correct word. The children read the passage silently and circled the words they believed were correct.

The results of administering these tasks are presented in Table 1. It is evident that the normal children were more proficient in oral reading than the poor readers since the mean of the former group was about 90 percent correct and the mean of the latter group was about 81 percent correct. This difference was statistically significant ($t = 2.10$, $df = 24$, $p < .05$). In addition the normal group was more proficient in comprehension since they attained a level of 89 percent accuracy on the maze task whereas the poor group attained a level of only 61 percent accuracy. This difference was also statistically significant ($t = 3.74$, $df = 24$, $p < .005$). Although the normal and poor groups were matched on standardized sight vocabulary and IQ, the latter group was slightly inferior in oral reading, about 10 percentage points, and substantially inferior in comprehension, about 30 percentage points.

The question of whether some children in the two groups are deficient in comprehension while other children are normal may be raised. To address this issue, a ratio of comprehension level and oral reading level was established for each child. For each subject, a ratio was constructed of the percent of the items correct on the maze task as the numerator and the percent of the words correct on the oral reading task as the denominator. Although the scales for the maze and oral reading tasks are not identical, a ratio less than 1.00 suggests that the child read orally material that he did not comprehend and a ratio greater than 1.00 suggests that the child's comprehension surpassed his oral reading.

A display of the ratios for both normal and poor readers is presented in Figure 1. It is evident that the ratios for a majority of the normal children clustered around 1.00. The mean was .99 with a standard deviation of .09. This indicates that the normal children understood the sentences very nearly to the extent that they could read them orally. A child who orally read 70 percent of the words correctly was likely to perform at the 70 percent level on the maze task. The poor readers, however, had a lower general level and were more different from each other than the normals. The mean of the ratios for the poor readers was .72 and the standard deviation was .21. This mean was significantly lower than normal ($t = 4.50$, $df = 25$, $p < .01$) and the standard deviation was significantly larger for the poor group than the normals ($F = 5.50$, $df = 12/12$, $p < .01$). From Figure 1 it can be observed that some of the poor group had ratios of about 1.00 which shows they could comprehend as well as they could read aloud. Other poor readers, however, had ratios in the .50 vicinity which shows that their comprehension was deficient in comparison to their oral reading level. These children need particular emphasis placed on instruction in comprehension in contrast to phonics, sight vocabulary or written expression.

An interesting relationship between comprehension and oral reading emerges from the correlation of the maze/oral reading ratios and oral reading level for both groups. The correlation for normal children was .26, whereas the correlation for the poor readers was .82. Among normals, the ratio was about 1.00 regardless of whether the oral reading was 50 percent or 100 percent. They understood what they could read orally and not more or less. For poor readers, children with high oral reading (90 percent) had high ratios (90 percent). But children with moderate oral reading (70 percent) had low ratios (50 percent). A vivid restatement of this relationship is that for poor readers whose oral reading was 90 percent, the maze accuracy was 90 percent. But for those whose oral reading was 70 percent, the maze was an astonishing 30 percent. In other words, when normal children have difficulty reading orally, some comprehension of the passage is possible. However, when poor readers have difficulty with oral reading, comprehension is drastically reduced. It appears that for poor readers the reading comprehension mechanism is delicate and breaks down rapidly when input from the decoding mechanism is even slightly distorted or reduced.

In summary, the results of the experiment were that poor readers were worse than normal readers who were matched with them on sight vocabulary. The poor readers were inferior on both oral reading and comprehension. The inferiority of the poor readers in comprehension was more marked than their difficulty in oral reading. Furthermore, normal children who could not read aloud with high proficiency understood at least some of the material. However, the poor readers who made even a few errors while reading the words in the passage did not comprehend the material at the most elementary level. It is clear that assessing, teaching and monitoring reading comprehension is particularly important for poor readers.

How can teachers, reading specialists and learning disability specialists use the maze technique? First, it can be used to identify children who have comprehension problems. A passage of 100 words from a basal story may be rewritten in maze form and given to a child. The child simply circles the correct alternatives and returns the material to be corrected by the teacher. The teacher then asks the child to read the words from the passage (either in isolation or in context) and the correct number of words is recorded. The percent correct on the maze may then be compared to the percent correct on the oral reading task. If the maze percentage is drastically lower than the oral reading percentage, (20 percentage points or more) comprehension is a distinct problem for the child. Teaching efforts can then be directed to the child's comprehension deficiency.

The maze technique may be used to monitor progress in comprehension. To accomplish this, the teacher may rewrite different basal stories into maze form and administer them periodically to determine a child's growth in comprehension. The maze can also be given in groups or subgroups at different levels. A chart such as the one represented in Table 2 may be used. The necessary information includes only: the child's name, date of administration of the maze task, book level of the passage, and the child's accuracy in terms of percent correct. If the maze task is given weekly and plotted over a 4 - 6 week period, the child's growth or lack of growth will be clearly observable to both the teacher and child.

From the chart in Table 2 it is clear that the child's comprehension at the 2^1 level was about 60 percent proficiency on September 12 - 14. His performance improved until it was over 85 percent for 3 different days, the 26 - 28th. Then material at the 2^2 level was introduced which was more difficult resulting in 75 - 80 percent proficiency. By Oct. 12 his performance was not above the 85 percent level and more teaching at that level would be necessary for this child.

The basic principles for using the maze procedure to guide the selection of materials are outlined below. Passages used in each maze should be new to the child and vocabulary for the passage should not be presented. Passages should be about 120 words long. Approximately every fifth word or less should be replaced with 3 alternatives. Alternatives should include: 1) the correct word, 2) an incorrect word that is the same part of speech (noun, verb, modifier, function) as the correct word, and 3) an incorrect word that is a different part of speech. The positions of the words should be varied. If a child is performing at about 90 percent accuracy for 3 - 4 administrations of the maze, more difficult material should be introduced. Optimal teaching levels are about 60 - 70 percent accuracy. If the child performs on the 3^1 material at the 50 percent accuracy, 2^2 at 65 percent and 2^1 at 90 percent accuracy, materials at the 2^2 level should be used for teaching. Materials below the 2^1 level cannot be used as easily as materials at higher levels. When the child reaches 85 - 100 percent accuracy at a given level, say 2^2, the next higher level should be introduced. When he reaches 85 - 100 percent accuracy on that level, materials at a higher level may be used and this cyclic process continues.

Considerable experience in using the maze task to monitor reading comprehension has been gained by 4 teachers in the Kennedy School in

Baltimore. During the past year 14 poor readers aged about 11 and reading at the 2.0 - 3.0 grade level were given a comprehension-oriented curriculum in reading. The children's deficiencies were primarily related to semantic and syntactic development rather than visual-perceptual skills. Maze tasks were used to guide the comprehension instruction and the following observations emerged from the teacher's experience. At the outset, children at the 2.5 grade level were given 2^1 readers. Performance at this level increased from 60 to 90 percent accuracy in about 2 weeks as the children learned the basic strategies for performing the task. When consistant mastery (85 - 100% accuracy) was reached at the 2^1 level, 2^2 level materials were introduced. This shift, or any increase in difficulty, may produce: 1) no change in comprehension, 2) fluxation in comprehension (65 - 85) percent accuracy on different days), or 3) reduction in comprehension (65 - 70 percent accuracy over several days). Often a child experienced a 3 week period of fluxation in accuracy (65 - 85%) before consistent performance over 85 percent at that level was attained.

Several factors were found which produced unusually low performance (below 60%) on the maze task, during the course of instruction in comprehension. Occasionally a story was extremely difficult due to uncommon vocabulary or proper names. If a child was emotionally distraught or fatigued from adverse home situations, his reading comprehension was temporarily decreased. Attempts to improve the speed of work usually degraded the accuracy of maze task performance and the inconsistent use of materials that varied 1 or 2 years in difficulty appeared to be confusing and caused decreases in comprehension.

During this teaching period, the technique provided a sensitive indication of comprehension and suggested when to increase or decrease the difficulty of materials to maintain growth at the fastest possible rate. In the 4 month period when the maze-guided instruction occurred, the average gain was 2 years in reading comprehension on the maze task. The majority of children gained 2 years in comprehension while a few gained 3 years and a few gained 1 year as determined by maze task performance.

The maze procedure has been presented as an assessment and monitoring technique, not a teaching method. A multitude of instructional activities and programs may improve comprehension. These improvements will be reflected in maze scores. Teaching word meanings, sentence structure, listening skills and phonics will likely succeed in increasing maze scores. The salutary value of using the maze procedure is that it serves as a thermostat to regulate comprehension instruction. It is not intended to serve as the furnace to raise the comprehension levels. The creativity and initiative of the teacher are needed to fulfill the essential teaching role.

Charting performance with the maze technique can be motivating for the student and the teacher alike especially if the children draw their own charts. Children can compete with themselves instead of their peers and can take deserved pride in their progress when it is undeniably reflected on the chart. Charts may be used by teachers to determine whether the teaching approach is effective and to illustrate the child's growth to others, such as parents or principals, who may wish to know about the child.

REFERENCES

Guthrie, J. T. Reading Comprehension and Syntactic Responses in Good and Poor Readers. Journal of Educational Psychology, 1973, 65 (3), 294 - 300.

Oldridge, O. A. A Congruent Validity Study of the Wide Range Achievement Test at Grade Seven. Educational and Psychological Measurement, 1964, 69, 107 - 113.

Washington, E. D. and Teska, J. A. Relations between the Wide Range Achievement Test and the California Achievement Test, the Stanford-Binet and the Illinois Test of Psycholinguistic Abilities. Psychological Reports, 1970, 26 (1), 291 - 294.

TABLE I

Oral Reading and Comprehension in the Maze Task

Subjects	Tasks		
	Oral Reading	Maze Performance	Maze/Oral Ratio
Normal $\overline{X}$	89.54	89.23	.99
SD	7.11	11.81	.09
Disabled $\overline{X}$	80.92	60.62	.72
SD	12.91	24.88	.21

Note: Scores in the maze/oral column are ratios.
Other scores are percent correct.

Table 2

Maze Chart

Name: ____________________

Teacher: ____________________

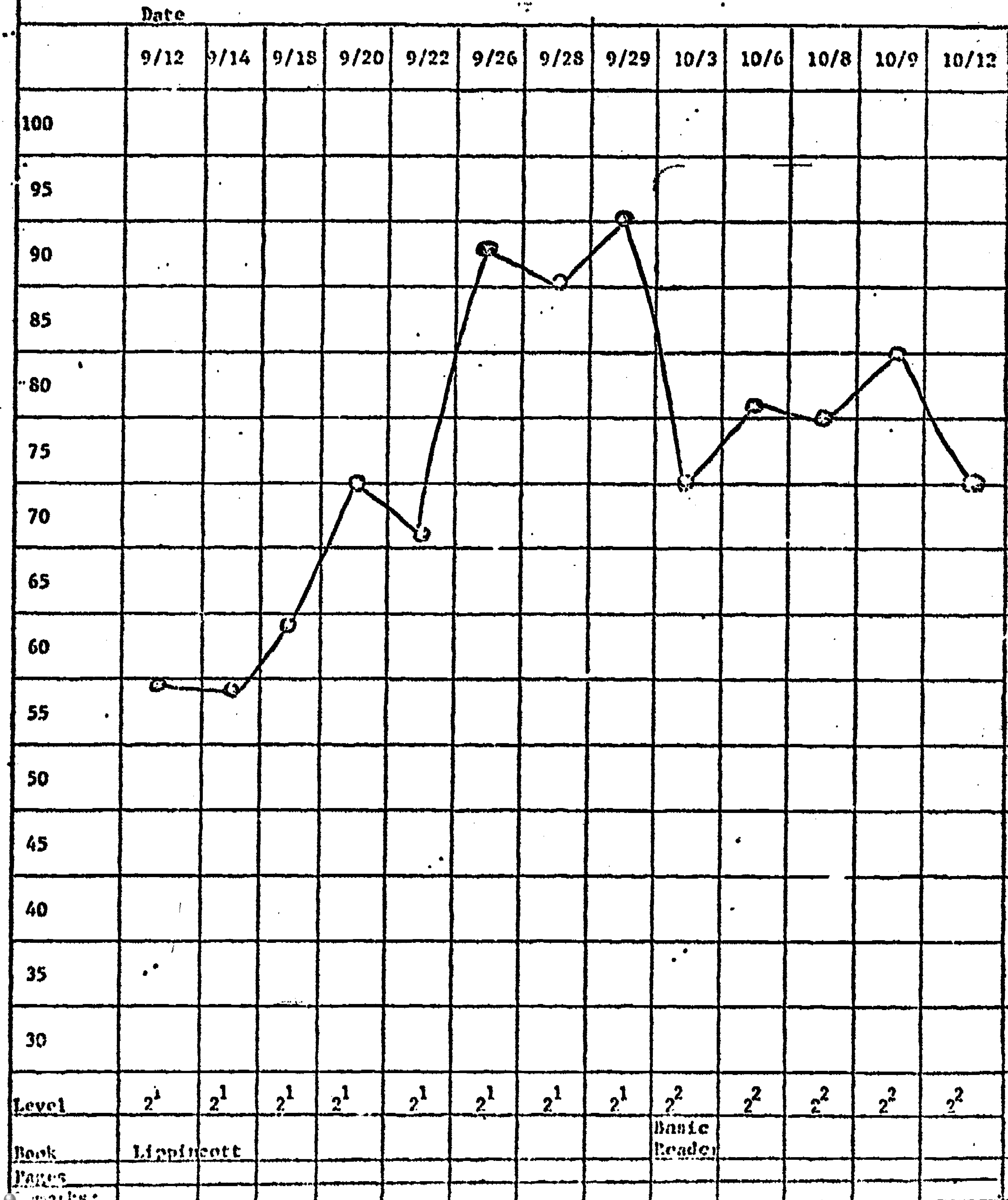
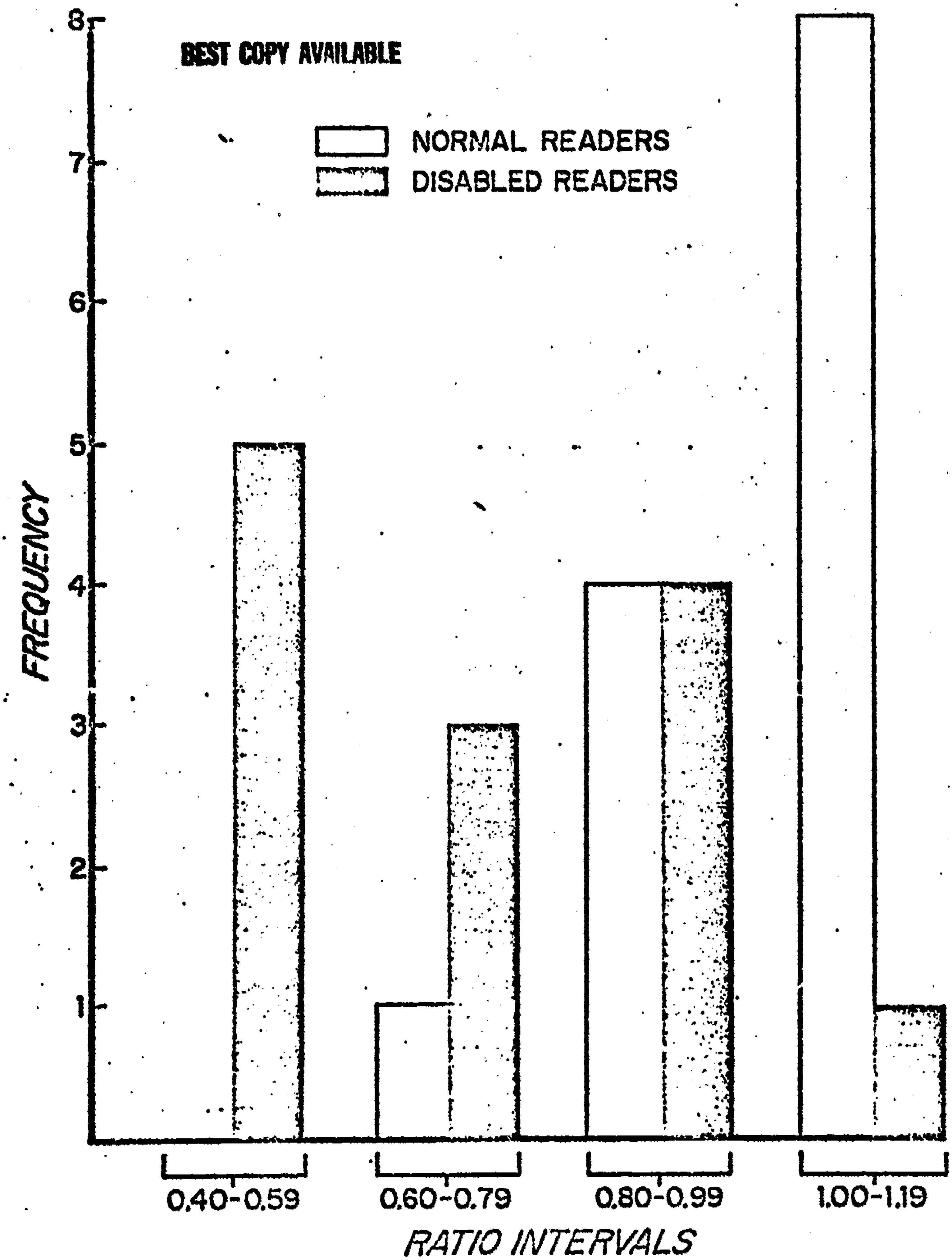


FIGURE CAPTION

Figure 1 Distribution of maze/oral reading reading ratios in normal and disabled readers.


NORMAL READERS
DISABLED READERS
FREQUENCY
8
7
6
5
4
3
2
1
0.40-0.59
0.60-0.79
0.80-0.99
1.00-1.19
RATIO INTERVALS

 rough draft

SCANNING TIMES THROUGH PROSE AND WORD STRINGS FOR VARIOUS TARGETS BY NORMAL AND DISABLED READERS

Rick Steinheiser and John T. Guthrie

Kennedy Institute, Johns Hopkins University

A group of disabled readers and a group of age-matched and a group of reading-level matched readers participated in a target scanning task. While reading a simple five sentence passage, the subject searched for a certain type of target: words of a given category, the letter "o", or the phoneme /ae/. All three groups were fastest in finding the word targets and slowest in finding the phoneme. The disabled readers were slower than the age-matched group for all targets. It was concluded that the phonological coding occurred at different rates after graphemic and semantic coding for the three groups.

SCANNING TIMES THROUGH PROSE AND WORD STRINGS FOR VARIOUS TARGETS BY NORMAL AND DISABLED READERS

Several recent word-scanning and letter-scanning studies of good and poor readers have failed to find any significant differences between the good and the poor readers (Katz & Wicklund, 1971; Katz & Wicklund, 1972). Assuming that a word can be encoded in terms of graphemic, phonemic, and semantic features, it should be possible to measure the relative encoding times for each of these sets of features (Cohen, 1970). The purpose of the present experiment was to determine if poor readers were significantly slower with any of these encoding steps than normal readers.

If the order of encoding of words is from grapheme to phoneme (and perhaps to spelling patterns) and then to "meaning," then one might expect search times to be shortest for a letter target, longer for a phoneme target, and longest for a semantic target. This kind of sequential processing has been reviewed by Neisser (1967). Alternative models (Smith, 1973) and data (Reicher, 1969; Wheeler, 1970; Smith & Haviland, 1971) raise the possibility that encoding may proceed from a fairly high level of abstraction ("meaning") of the stimulus to the analysis and synthesis of lower-order units such as phonemes (phonetic representation). This would imply that words are perceived as synthesized entities, from which an awareness of the constituents is subsequently derived. That is, the synthesis process utilizes fragments of the graphemic-phonemic matrix, but these fragments are not accessible for volitional processing (recognition) until the larger unit (word) has been constructed.

Thus, poor readers could be inferior to normals (as measured by total scanning time through passages) in terms of ability to find a visual target (a specific letter), a phonemic target (a short vowel sound), or a word (of a specified semantic category). Dependency upon syntactic cues may also be assessed by examining performance on sentences in which the word order has been randomized. Presumably normal readers would show a greater proportional increase in search times in this condition over the normal sentence structures than the disabled readers. This prediction was made because the reading speed of the disabled readers might be so slow with the normal sentences so as to render useless any syntactic information which facilitates high-speed scanning (Gough, 1972).

<u>Subjects</u>. Three groups of twelve subjects each were used. Each of these subjects had participated in previous reading experiments (Steinheiser & Guthrie, 1973). The disabled group had a mean chronological age of 10 years 6 months, a median I.Q. of 98 (on the WISC-FS), and a median vocabulary of grade 2.8 (on the Gates-MacGinnitie). Group II students had a mean age of 7 years 4 months, median I.Q. of 105 (on the Peabody Picture Vocabulary Test), and a median vocabulary of grade 2.6. Group III students were from the same parochial school as Group II, and they had a mean age of 10 years 3½ months, a median I.Q.

of 105 (on the PPVT), and a median vocabulary of grade 6.8.

Tests for statistical significance showed that the disabled group was older than the reading-level matched group ($t=1.98$, $df = 34$, $p < .06$); and the age-matched control group was also older than the reading-level matched group ($t=2.16$, $df = 34$, $p < .05$). An extension of the median test produced a significant chi-square ($X^2 = 22.66$, $df = 2$, $p < .01$) from the vocabulary scores of the three groups. (All of the vocabulary scores for the age-matched group exceeded the grand median.) The I.Q. scores of the three groups yielded an insignificant 4.3 on the extension of the median test.

Stimuli. Six short stories were used, ranging from 23 to 30 words per story and 4 to 5 sentences per story. Each story was typewritten on a sheet of paper for inspection by the subject. Each story was about a different topic, e.g., foods, clothes, animals. Each story contained on the average five "targets." The five semantic targets varied from story to story; for example, cow, cat, dog, fox, and pig were the animal targets in one story, and blue, green, red, brown and black were the color targets in another story. Each story also contained an average of five instances of the letter "o" and five instances of the short "a" phoneme /æ/ as in "ran" and "cat."

In one condition the sentences were unaltered, whereas in another condition the words within a given sentence were scrambled, but the order of occurrence of each scrambled sentence within the passage itself was maintained.

Procedure. Half of the subjects from each group were presented the normal sentences on the first day of testing and the randomized set on the second day. Counterbalancing for the three types of target was also performed.

The first task of each subject was to read each of the six stories aloud. This was done to assure that the selection of words was within the pronunciation range of the subject.

The next task for all subjects was to read each story as fast as possible and mark the period at the end of each sentence in that story. Most subjects elected to read silently, although several subjects in the younger and remedial groups tended to read aloud (whispering to themselves). The experimenter was able to assure that each reader actually did read each sentence (instead of simply scanning for the period) by observing the reading behavior (eye movements, movement of the index finger along the words of a sentence) of the subject. The reader was told that he would be timed in this and subsequent tasks by the experimenter who held a manual stopwatch accurate to the nearest .1 second. The reading time (to mark the periods) for each story for each subject was recorded.

In the short "a" vowel (/æ/) search, the subject was instructed to mark all of the letter a's in the story which had a "short" sound, e.g., cat, ran, and.

In the letter search task, the subject was given a fresh page of the same stories, and asked to mark all of the letter o's contained in all of the words in each story.

In the semantic category search the reader was given a new sheet with the six stories and was asked to mark all of the words of a given story which were instances of a particular category, e.g., all of the animal names in one story, all of the color names in another story, etc. The reader was given feedback (search time, and any errors that were made) in each of the target searches.

The first story of each set was always considered practice, and was omitted from data analysis. The main data of interest were the search times for each reader to the five targets in the five different stories for the four different types of targets.

In this split-plot design, the three groups of subjects was the between variable, and the type of sentences (normal vs. scrambled) and the four types of targets were the within variables.

Results

Figure 1 shows the average search times for each group across the five normal test stories, and Figure 2 shows the search times for the same targets within scrambled sentences. Analysis of variance of these data is presented in Table 1. It may be seen that all of the main effects (groups, sentence type, and target) are significant by the Geisser-Greenhouse test with reduced degrees of freedom.

Inspection of Figures 1 and 2 shows that there was a minor effect due to sentence structure for the remedial and age-matched groups, but that the younger group was consistently much slower in detecting all types of target in the scrambled sentence condition. A Scheffe test (see Table 2) showed this to be a highly significant difference ($F=21.86$, $p<.01$).

The general conclusions that can be drawn from the comparisons in Table 2 are these: the disabled readers were faster with the scrambled sentences than the second-graders; the second graders were faster on the sentences than with the scrambled versions; all three groups were generally fastest with word targets, slower to the letter target, and slowest when searching for the phoneme target. Note that the relative difficulty for the phoneme target was extremely great for the disabled readers as compared to the other groups and other targets.

The disabled group was consistently slower to all targets in both conditions than the age-matched group. The greatest difference between them occurred in the phoneme search. The disabled group was also faster than the reading-level matched group when searching for words and for a single letter. These two groups produced similar latencies to the phoneme target in both the sentence and scrambled sentence condition.

Discussion

All three groups were consistently faster in finding the semantic (word) target than the letter target; and the letter was consistently found faster than the phoneme. This finding is in general agreement with the results obtained by Cohen (1970). One interpretation of this finding is that it suggests the order of processing to be semantic (the whole word), then graphemic, and then phonological (Smith & Holmes, 1971).

A more cautious interpretation may be warranted, which arises from the possible discrepancy between the size of the potential target population in the various search conditions. Since there were fewer words than letters in each story, it might seem reasonable that words should be recognized faster than letters simply because there are fewer words to process. Although possible, this argument seems to beg the question, since word recognition may be either serial (in stages) or parallel (holistic, simultaneous processing of all constituents). To say that there were fewer word targets than letter targets per story is to ignore the constituents of a word, namely, letters. The "unequal target probability" hypothesis would treat words as separate holistic targets, and this becomes equivalent to the order of processing proposed by the authors to explain the results: semantic, graphemic, and phonological.

A developmental effect was clearly evident, in that the younger group's performance deteriorated significantly when syntactic cues were no longer available. This contradicts the hypothesis proposed earlier in the paper, but does suggest that the disabled readers were able to focus attention upon the task at hand better than the younger group. That is, the lack of linguistic structure may have served as an attentional test for the disabled readers, and less as an indicator of their psycholinguistic capabilities.

The present results confirm Cohen's (1970) findings: semantic feature search is faster than for a letter (graphic feature), and both were significantly faster than searching for a phonological feature. However, the results are somewhat disparate from those reported earlier by Katz and Wicklund (1970), since differences between the good and poor readers did emerge in this study. For example, in the present study, the disabled readers were considerably slower than the age-matched control group even when searching for the word (semantic) target. Possibly the disabled readers used in the present experiment were considerably more disabled than those used by Katz and Wicklund ("Subjects in each grade were tentatively divided into groups of good and poor readers based on the reading test scores which were available." Katz and Wicklund, 1972, p. 364).

Finally, the phonological target was consistently the one that required the longest search times. Since the disabled group took the longest time with this target relative to the other targets, it would seem that their source of difficulty comes at the stage in which "sight" is to be encoded into "sound." It appears that the word is encoded as a "gestalt," from which specific graphic information may then be extracted, along with the segmentation into pronouncable sequences. Rather than constructing a word from its constituent letters, it would appear that the word is the unit of analysis, from which more refined information (semantic, graphic, phonological) is segmented. Neisser

(1967) has called these hypotheses "inference" and "unitization." Results from the present experiment, and also those of Smith and Haviland (1972) tend to support the unitization hypothesis. Or, as Smith (1973) contends, "...meaning can be extracted from sequences of written words independently of their sounds, and that in fact meaning must be comprehended if sounds are to be appropriately produced (p. 78)."

REFERENCES

Cohen, G. Search times for combinations of visual, phonemic, and semantic targets in reading prose. Perception & Psychophysics, 1970, 8, 370-372.

Gough, P. One second of reading. In J. F. Kavanagh and I. G. Mattingly (Eds.): Language by ear and eye: the relationships between speech and reading. Cambridge, Mass: M.I.T. Press, 1972.

Katz, L. & Wicklund, D. Simple reaction time for good and poor readers in grades two and six. Perceptual and Motor Skills, 1971, 32, 270.

Katz, L., & Wicklund D. Word scanning rate for good and poor readers. Journal of Educational Psychology, 1971, 62, 138-140.

Katz, L., & Wicklund, D. Letter scanning rate for good and poor readers in grades two and six. Journal of Educational Psychology, 1972, 63, 363-367.

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

Reicher, G. M. Perceptual recognition as a function of the meaningfulness of the stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Smith, F. Psycholinguistics and reading. New York: Holt, Rinehart, and Winston, 1973.

Smith, E., & Haviland, S.E. Why words are perceived more accurately than non-words: Inference vs. unitization. Journal of Experimental Psychology, 1972, 92, 59-64.

Steinheiser, R., & Guthrie, J.T. Perceptual and linguistic processing of letters and words by normal and disabled readers. Submitted for publication, 1974.

Wheeler, D. Processes in word recognition. Cognitive Psychology, 1970, 1, 59-85.

TABLE 2

SCHEFFE POST-HOC COMPARISONS

Comparison	F	Probability
A1C2-A2C2	2.88	N.S.
A2C1-A2C2	21.86	.01
B2-B3 at A3C1	8.27	N.S.
B2-B4 at A3C1	15.00	.01
B2 B4 at A2C1	8.95	N.S.
B3-B4 at A1C1	43.02	.01
B4-B3 at A2C2	8.50	N.S.
B2-B4 at A3C2	13.50	.01
B3-B4 at A1C2	48.62	.01

Groups (A)

Key: A1 - Disabled Readers
A2 - Second Graders
A3 - Fourth, Fifth Graders

Targets (B)

B1 - Period
B2 - Word
B3 - Letter "o"
B4 - Phoneme /oe/

Structure (C)

C1 - Sentences
C2 - Scrambled Sentences

TABLE 1

THREE-WAY ANALYSIS OF VARIANCE

Source	d.f.	Mean Square	F	P	d.f. for F
A (groups)	2	3377.38	43.1	.001	2,33
subjects w. gps.	33	78.5			
B (target item)	3	1048.91	67.7	.001	1,33
AB	6	123.08	7.93	.01	2,33
B x subjects w. gps.	99	15.5			
C (structure)	1	299.72	16.2	.01	1,33
AC	2	71.83	3.87	.05	2,33
C x subjects w. gps.	33	18.57			
BC	3	21.32	20.1	.01	1,33
ABC	6	7.90	7.45	.01	2,33
BC x subj. w. gps.	99	1.06			

ACKNOWLEDGEMENTS

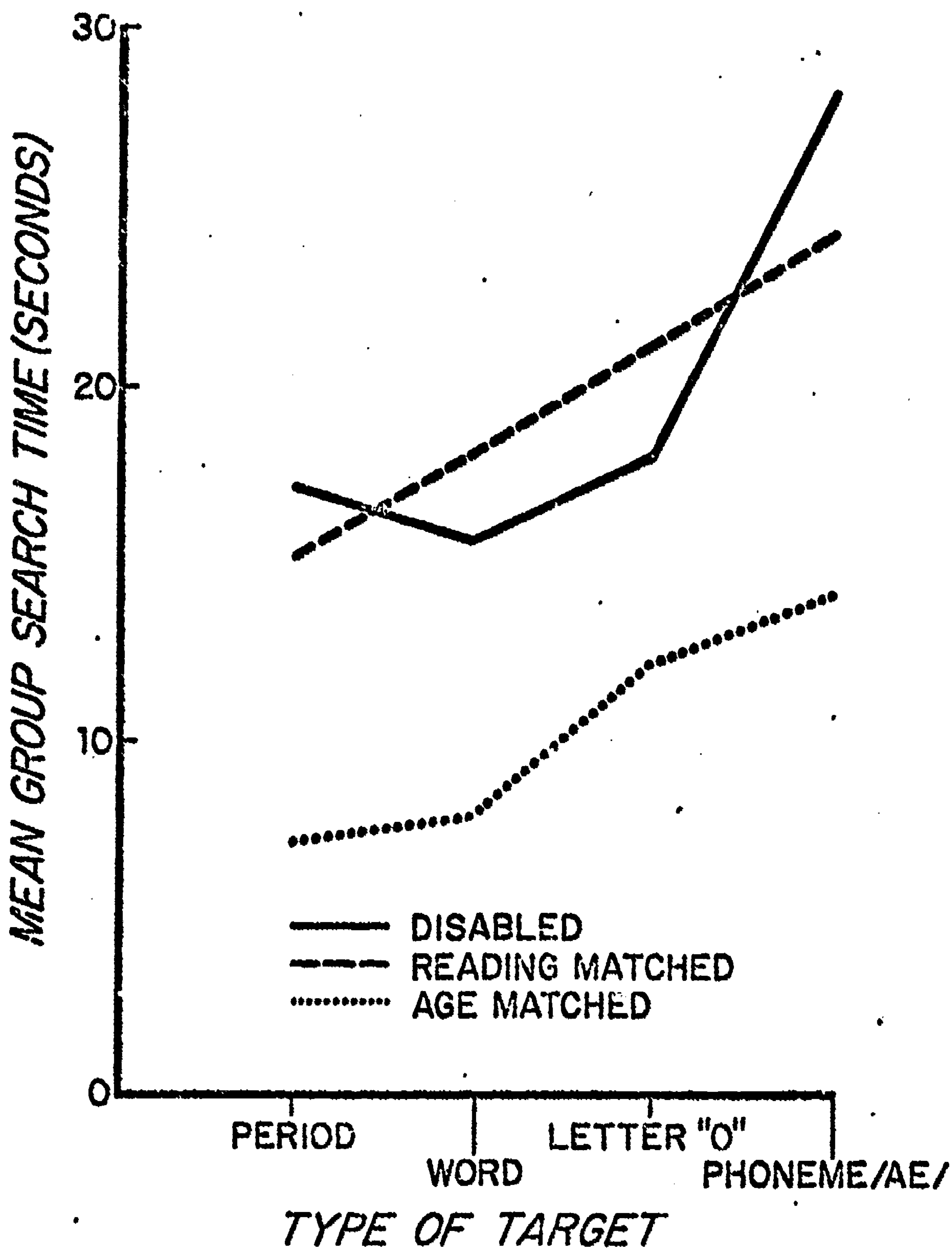
[1]Assistance in the conduct of this study was provided by the Principal and staff of the Kennedy School for Special Learning Disabilities and the Principal and staff of Our Lady of Victory School.

Support for this experiment was provided by Maternal and Child Health Project #917, the Joseph P. Kennedy, Jr. Foundation, and the Spencer Foundation.

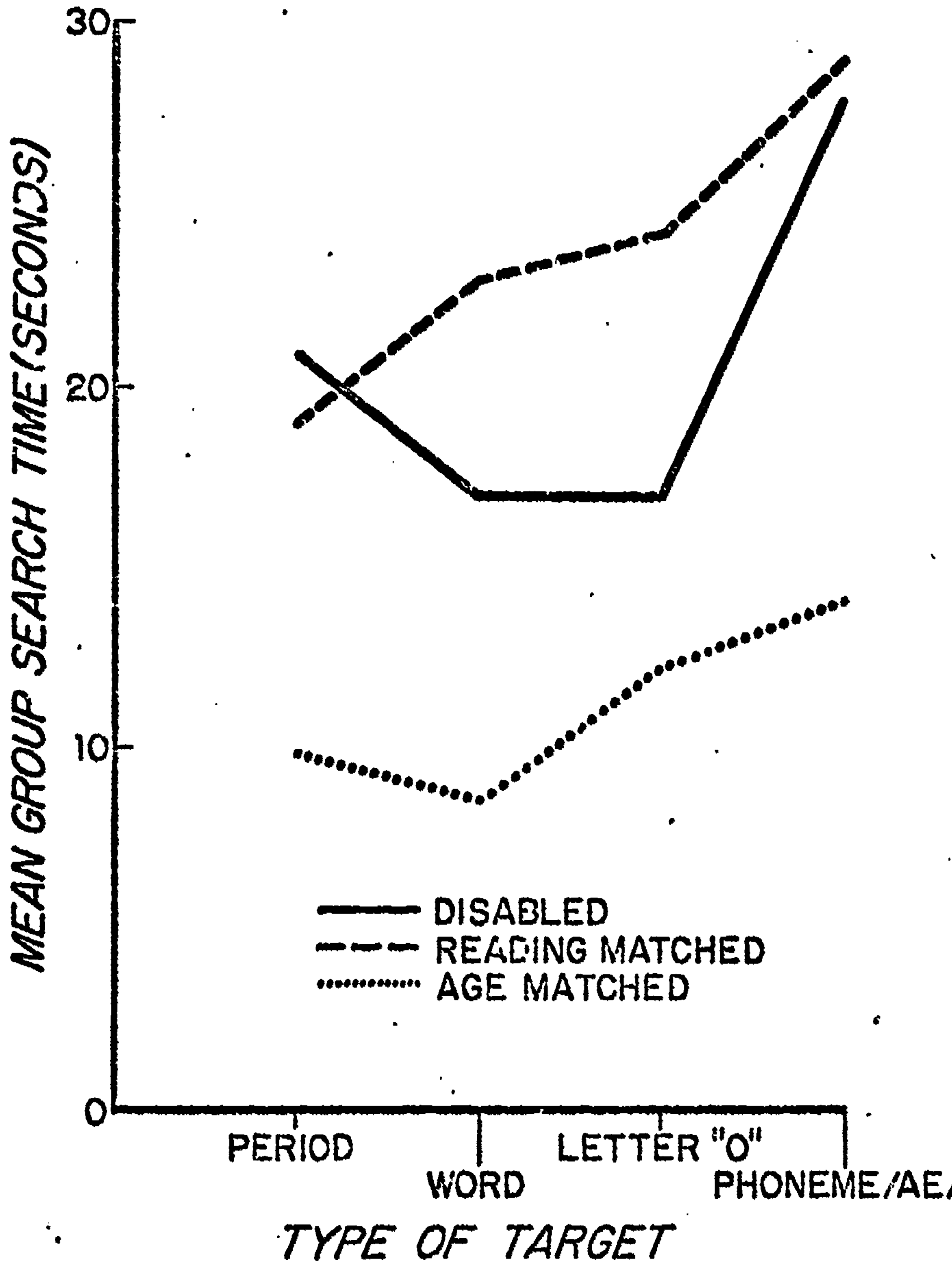
FIGURE CAPTIONS

Figure 1. Average group scanning times for target in stories with normal sentence structure.

Figure 2. Mean group scanning times for targets in stories with a scrambled sentence structure.



MEAN GROUP SEARCH TIME (SECONDS)
30
20
10
0
DISABLED
READING MATCHED
AGE MATCHED
PERIOD
WORD
LETTER "O"
PHONEME/AE/
TYPE OF TARGET


30
20
10
0
MEAN GROUP SEARCH TIME (SECONDS)
DISABLED
READING MATCHED
AGE MATCHED
PERIOD
WORD
LETTER "O"
PHONEME /AE/
TYPE OF TARGET

ADEQUATE VISUAL ENCODING FOLLOWED BY POOR PHONOLIGICAL ENCODING AS A SOURCE OF READING DISABILITY

Frederick H. Steinheiser, Jr. and John T. Guthrie[1]

Kennedy Institute, Johns Hopkins University

ADEQUATE VISUAL ENCODING FOLLOWED BY POOR PHONOLOGICAL ENCODING AS A SOURCE OF READING DISABILITY

Frederick H. Steinheiser, Jr., and John T. Guthrie

Abstract

A group of disabled readers, and a group of age-matched and a group of reading-level matched readers participated in a "same-different" reaction-time task. After viewing a pair of items (non-alphabetic symbols, 4-letter vowel strings, or real words), the subject indicated whether the same graphemic symbol had been underlined in each item of the pair. A second task required an additional phonological encoding step: the underlined letters of the two words could be different yet still produce roughly the same vowel sound when pronounced: e.g., heat-feet. Also, the same letters in the two words could produce different sounds: e.g., tough-shout.

In the visual comparison tasks, the only significant effect was that the younger normal group was slower than the age-matched control group. However, when phonological coding was required, the disabled group was significantly slower than the young normal readers, and they in turn were slower than the age-matched normals.

These results suggest that disabled readers have adequate visual processing abilities, but poor phonological encoding. One implication is that methods of reading remediation might concentrate upon spelling patterns, in order to exploit the systematicity between graphemes and phonemes.

ADEQUATE VISUAL ENCODING FOLLOWED BY POOR PHONOLOGICAL ENCODING AS A SOURCE OF READING DISABILITY

English orthography is reputed to be disorderly with respect to the phonetic realizations of written words. For example, the letters "th" are pronounced /θ/ in "thin" but as /ð/ in "this," and "ea" is pronounced as /i/ in "heat" but as /ɛ/ in "head." The list of such examples is extensive, and far from complete (Venezky, 1967). A common anecdote is that the person who learns to read English as a native language and then learns a more regular orthography such as Spanish or Finnish is far better off than the person who learns to read (and speak) English as a second language.

Since reading a given language must be explicitly taught to children (unlike their spontaneous learning to speak and comprehend spoken discourse), the problem arises as to how best to provide initial reading instruction. One approach (phonics) emphasizes the sight-sound (grapheme-phoneme) correspondence in printed and spoken words. This may include teaching the child to explicitly label grapheme clusters as one or another type of utterance; e.g., "cat" has a "short a" and "lake" has a "long a." A different approach stresses that the child shouldlearn that printed words are signals for spoken words, and that meaning can be apprehended from these printed words directly. Thus, the graphic symbols "c a t " stand for that small furry animal with a tail, claws on its feet, that many people have as a pet, etc., that we call "/kæt/" in spoken discourse.

Considerable research has been conducted to determine the nature of reading disability. The tasks have included measures of oral reading speed, scanning rate for target words in prose and word lists (Katz and Wicklund, 1972), and sentence completion (Guthrie, 1973 a), The sources of reading disability have been attributes to inadequate cognitive (linguistic) skills (Guthrie, 1973 b), poor mastery of grapheme-phoneme correspondence (Samuels, 1973), blocked transmission of impulses through the angular gyrus (Preston and Guthrie, 1973), and the very English orthographic system - as an example of all alphabetically based writing systems (Rozin, Poritsky, & Sotsky, 1971).

One purpose of the present experiment was to examine the "psychological reality" of certain spelling units commonly found in English words. Various word perception studies (Reicher, 1969; Wheeler, 1970; Gibson, 1971; Smith & Holmes, 1972) have shown that complex internal relationships among the constituents (distinctive features, graphemes, spelling patterns) interact with the subject's expectancies (Aderman & Smith, 1971). There are many processing stages involved in correctly perceiving a word, one of which involves grapheme-phoneme correspondence rules. The psychological reality of this rule system was explored in this experiment, in which both normal and disabled readers participated. By explicitly classifying the grapheme-phoneme relationship in a set of words, it was hoped that the disabled readers' performance with these categorized words would deviate in a consistent manner from that of the normal readers.

Venezky (1967) has developed an elaborate system of rules which purports to logically characterize the grapheme-phoneme relationship in many English words. As an example, consider the two letters "ea." When these letters occur in this adjacent order in a word, the phonetic realization of the vowel depends upon the context provided by the surrounding letters. Thus, "ea" is pronounced / ĭ / in "bead" and as /ɛ / in "head." There is a one-to-many mapping from the "ea" graphemes to the possible / ĭ / or /ɛ / phonemes. The list of similar examples is extensive: /ou/ (could-count), /th/ (this-thin), /a/ (cat-lake), /u/ (put-cup), /i/ (hit-like), /ai/ (rain-said), etc.

In the present experiment, a simple 2 x 2 grapheme-phoneme matrix was used to classify the words. In Fig. 1, quadrant 1 contains grapheme (ea) which correspond to very nearly the same vowel phoneme / i /, in such words as heat and bead; quadrant 2 words contain different graphemes which produce highly similar phonemes (e.g., "ee" and "ea" in the words heat and feet both produce the phoneme / ɪ /; quadrant 3 words contain the same relevant graphemes but produce different sounding phonemes (e.g., "ea" is in both heat and head, but the vowels are different); and quadrant 4 words contain different relevant graphemes which produce different phonemes (e.g., heat and said). The experimental question was: will normal and disabled readers be able to classify word pairs from each quadrant more easily on the basis of "same-different graphemes" than on the basis of "same-different phonemes?"

Presumably the additional coding stage in going from grapheme to phoneme would increase the response latency for all subjects. That is, it should take longer to decide that "heat" and "head" contain different vowel phonemes than to decide that they both contain the same (ea) middle graphemes. Comparison of the disabled readers' performance on the two tasks with that of normal readers should indicate whether their source of difficulty is visual perceptual or in the phonetic encoding.

Subjects. Three groups of elementary school children participated, each group comprising 12 students. Group I students attended a remedial reading program at the Kennedy Institute in Baltimore. This group had a mean chronological age (CA) of 10 years 6 months, a median I.Q. of 98 (on the WISC-FS), and a median vocabulary of grade 2.8 (on the Gates-MacGinitieTest). Group II students attended a parochial school in Baltimore, and were all in the second grade. The mean (CA) of this group was 7 year 4 months, with a median I.A. of 105 (on the Peabody Picture Vocabulary Test), and a median vocabulary of 2.6. Group III students were from the same parochial school, and were in the fourth, fifth, and sixth grades, the mean (CA) of this group was 10 years 3½ months, with a median PPVT I.Q. of 105, and a median vocabulary of grade 6.8.

Tests for statistical significance showed that the disabled group was older than the reading-level matched group ($t = 1.98$, $df = 34$, $p < .06$); and the age-matched control group was also older than the reading level matched group ($t = 2.16$, $df = 34$, $p < .05$). An extension of the median test (Siegel, 1956) produced a significant chi-square

(X^2 = 22.66, df = 2, $p < .01$) from vocabulary scores of the three groups. (All of the vocabulary scores for the age-matched group were above the grand median.) The I.Q. scores yielded an insignificant 4.3 on the Extension of the Median Test.

Thus, groups I and II were approximately matched on reading level but not on age; whereas groups I and III were approximately matched on age but not on reading level.

Stimuli. Common monosyllabic words were selected which conformed to a simple 2 x 2 grapheme-phoneme matrix. In the present experiment, only the vowel of each word was of interest. Some examples will help to clarify the utility of using this matrix scheme for selecting stimuli. Both could and would have the same middle graphemes (ou) and the vowel sound is the same for both words; could and wood have different middle letters but the vowel sound is the same; could and count have the same middle letters but the vowel sound is different; and count and wood have both different middle letters and different vowel sounds. This matrix was used to generate a total set of 76 word pairs. Representative pairs from quadrant 1 are: heat-seat, could-would, put-push; from quadrant 2: heat-meet; could-book, put-look; from quadrant 3: heat-head, could-count, put-cup; from quadrant 4: heat-said, could-brown, put-cow. The 76 word pairs used in this study are merely representative of a much larger set of possible pairs, if less common and multisyllabic words were included. The critical letters were always underlined when each pair was presented to the subject.

Two sets of control stimuli were also used. One set consisted of the vowels a,e,o,u, and the other set consisted of the symbols @,#,&,*. Permutations of the 4 elements generated the pairs for same-different comparisons, with the constraint that only the two middle elements could be in different positions within each pair. Thus, representative pairs were: aeou-aeou, oacu-oaeu, uaeo-ueao, @#&*-@#&*, *&#@-*#&@. Sixteen same and sixteen different pairs from each control set were used.

Apparatus. Each pair of items was set in lower case type, a transparency made, and mounted in a 2"x2" slide for presentation by a carousel slide projector. The members of each pair were presented one above the other. Two push-buttons were provided for the S. A standard clock was synchronized to start when the tachistoscopic shutter in front of the projector opened, and to stop when S pushed either button. Button position was counter-balanced across Ss.

Procedure. Each S participated in three separate sessions, the order of stimulus presentation being counter-balanced across Ss. Before each session, the S was shown the pertinent list of test items on a sheet of paper, and was required to respond to each pair as being either the "same" or "different." In the two conditions when real words were used, S was also required to read each word aloud before deciding "same" or "different." E gave assistance or corrections when necessary.

In the letter comparison task, S was shown the list of 76 word pairs, and had to put the letter "s" (for "same") next to pairs from quadrants one and three, and a "d" (for "different") next to pairs from quadrants two and four. The same procedure was followed in the vowel comparison task, except "s" was to be placed before quadrant one and two pairs, and "d" next to quadrant three and four pairs. For practice with the control lists of non-words, S simply placed the "s" or "d" next to each pair.

After this familiarization phase, a set of 12 practice trials with the slides was taken. Feedback was always given during practice, and throughout the experiment. Trials on which a mistake was made were taken over again at the end of the list. If a mistake was made again, the response time to that pair was not included in the latency analysis. The longest latency for each S from each condition and each stimulus classification was omitted in order to roughly correct for momentary lapses of attention and to reduce the skewness of the latency distribution.

Experimental Design. In order to assess the effect of the task variable (same letters vs. same vowels), the three groups of Ss was the between factor, and two levels of instruction and four quadrants of the spelling-sound matrix were the within-subject factors in a "split-plot" factorial analysis of variance.

In the analysis of the control condition, the latencies to the same spelling quadrants (1 and 3) were grouped together, and another grouping from quadrants 2 and 4 (different spelling quadrants). The same-different latencies to each of the two sets of control pairs were included in this analysis. Again, the between-subjects factor was the three groups of children, while the within-subjects factors were type of response (same or different) and type of stimulus pair (words, letters, or non-alphabetic symbols).

Results

All main effects and interactions except the triple interaction attained some level of statistical significance. Table 1 presents the complete analysis of variance of the data.

The mean response latencies for each group in each condition are presented in Figures 1 and 2. The effect of the instructional variable is readily apparent, since the response latencies doubled (and in some cases almost tripled) in the phonemic matching condition over what they were in the letter matching condition. " isabled readers' performance was approximately midway between the control groups (1.44 vs. .98 and 1.69 sec.) in the letter identity condition. However, when vowel identity was the criterion for making a "same" response, the average of the disabled readers was 4.61 sec., vs. 3.09 for the younger group, and 1.76 for the age-matched group.

A set of Scheffe post-hoc tests was performed on various sets of means to determine exactly where the significant effects occurred. The only difference between groups in Fig. 1, that attained statistical significance occurred between the two control groups ($F = 8.71$; $df = 2.33$; $p < .05$). No comparison between any two stimulus quadrants attained significant or beyond the .05 level. The difference between the two conditions of instruction (an overall average of 1.37 sec. in Fig. 1 and 3.15 sec. in Fig. 2 was highly significant ($F = 32.02$; $df = 1.33$; $p < .001$).

As is visually apparent from inspecting Fig. 2, the average latencies of the three groups differ significantly from each other. Table 2 shows the results of these tes. Four out of the six possible comparisons among the stimulus quadrants also significantly affected response times. Stimuli from quadrants 1 and 4 produced the shortest latencies, while items from quadrants 2 and 3 produced the longest latencies. It is worth emphasizing that the younger control group produced the longest latencies in the visual match condition (Fig. 1), and that the disabled readers produced the longest latencies in the phonemic match condition (Fig. 2).

Results from the Control Condition

Analysis of variance yielded significant effects due to groups ($F = 20.07$; $df = 2,33$; $p < .001$), type of stimulus item ($F = 6.52$; $df = 1,33$; $p < .05$) and group x item interaction ($F = 7.96$; $df = 2,33$; $p < .01$). The df for the F ratios have been computed according to the conservative Geisser-Greenhouse method. No other effects reached statistical significance even when computed with uncorrected df. Inspection of Fig. 3 shows that the younger group was consistently slower than the disabled readers with all three types of items. Since the mean latencies of "same" and "different" responses did not statistically differ, they have been grouped together -- hence the absence of "same-different" categories in Fig. 2.

Scheffé tests were then conducted on various combinations of means to assess where the significant effects occurred. These are presented in Table 3. The differences between the three groups can be seen to be significant, except that the disabled readers did not differ significantly from the age-matched control group (A1-A3 comparison). The younger group was slower than the disabled readers (A1-A2) and much slower than the older control group (A2-A3). Latencies tended to be longest for the letter pairs (B1), shorter for the non-letter symbol pairs (B2), and shortest for the word pairs (B3). The young control group was the only group to respond significantly faster to one set of stimuli, being faster to real words than to letters or to non-letter symbols (B3 vs. B1 and B2 at A2).

Discussion. When the readers were able to use visual identity as the criterion for sameness (e.g., "heat" and "head" are the same because the same two letters were underlined in each word), the disabled readers performed faster than the reading level matched second graders, but more slowly than the age-matched 4th and 5th graders. However, the only significant difference occurred between the two control groups. The average latencies for the disabled readers did not differ significantly from either control group. This finding suggests that the disabled readers were able to make graphemic comparisons slightly faster than younger readers and slightly slower than age-matched normal readers. The 2 x 2 grapheme-phoneme correspondence matrix had no significant effect upon any group in this task.

When the same word pairs were presented in the phoneme comparison condition, the disabled readers were significantly slower than either control group. Further, the younger control group was significantly slower than the age-matched control group. This finding suggests that the two tasks have isolated reading disability as being qualitatively different from any simple form of developmental retardation. That is, in the visual task, the younger readers were slightly slower than the disabled group, but in the phoneme task, a complete and remarkable consistent reversal occurred. Fig. 2 shows that the phoneme task was indeed more difficult than the visual task of Fig. 1 for all 3 groups. But the former is proportionately more difficult for the disabled readers than for either of the control groups.

If we take the difference between the average latencies of each group for each task, we arrive at a rough index of grapheme-phoneme encoding efficiency. Thus, the older normal readers took an average of .98 sec. across the 4 item pairs in the grapheme matching task, and 1.76 sec. in the phoneme comparison task. This difference is .78 sec. Dividing this difference by the larger latency gives .44 (.78/1.76 = .44), indicating that the criterion change from grapheme to phoneme comparison nearly doubled the average response latencies for this group. For the younger readers, this difference is 1.55 sec., and the index is 1.55/3.09 = .50. For the disabled readers, the difference is 2.92 sec., with an index of 2.92/4.61 = .63. This substantiates the claim that the phoneme comparison task was proportionately more difficult for the disabled group.

The quadrant from which the stimulus pair was selected seems to have had little effect upon the older normal readers. What significant effects did occur from this variable arose mainly with the disabled and younger groups. It is reasonable that stimuli from quadrants 1 and 4 led to the shortest latencies, because a pair from quadrant 1 is physically the same, and a pair from quadrant 4 is physically different. However, pairs from quadrants 2 and 3 have no simple relationship between sight and sound, and hence took longer. It is noteworthy that even the older normal group took average of nearly twice as long to respond to the same items in the phoneme as in the visual comparison condition. Thus, physical identity and physical difference were not able to facilitate performance in the phoneme comparison task.

The results of the present experiement help to clarify some issues in the recent reading research. Both Guthrie (1973) and Posner, Lewis, and Conrad (1972) have emphasized the importance of individual differences in any taxonomy of reading subskills. In view of the present results, it seems that the subskill of encoding visually presented words into internally represented sounds is an extremely time-consuming process for all groups of readers, but significantly moreso for the disabled group. Thus, a prime source in reading difficulty may be the complex, but categorizable, relationships that exist between spelling and sound (Shankweiler & Liberman, 1972). The fact that the disabled readers showed such poor peformance in the sound-matching condition, even though the word pairs were always presented <u>simultaneously</u>, casts some doubt upon the suggestion recently proposed by Mackworth and Mackworth (1974): "... poor readers in grade 10 have a poor <u>long-term</u> visual memory for words... (p. 60, italics added)."

References

Aderman, L., & Smith, E.E. Expectancy as a determinant of functional units in perceptual recognition. Cognitive Psychology, 1971, 2, 117-129.

Gibson, E.J. Perceptual learning and the theory of word perception. Cognitive Psychology, 1971, 2, 351-368

Guthrie, J.T. Reading comprehension and syntactic responses in good and poor readers. Journal of Educational Psychology, 1973, 65, (3), 294-300.

Guthrie, J.T. Models of reading and reading disability. Journal of Educational Psychology, 1973, 65, 9-18.

Mackworth, J.F. & Mackworth, N.H. Spelling recognition and coding by poor readers. Bulletin of the Psychonomic Society, 1974, 3, 59-60.

Posner, M.I., Lewis, J.L., & Conrad, C. Component processes in reading: A performance analysis. In J.F. Kavanaugh and I.G. Mattingly (Eds.), Language by ear and by eye. Cambridge, Mass: M.I.T. Press, 1972.

Preston, M., & Guthrie, J.T., and Childs, B. Visual Evoked Responses (VER) in Normal and Disabled Readers. Psychophysiology, 1974, in press.

Reicher, G.M. Perceptual recognition as a function of the meaningfulness of the stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Rozin, P., Poritsky, S., & Sotsky, R. American children with reading problems can easily learn to read English represented by Chinese characters. Science, 1971, 171, 1264-1267.

Samuels, S.J. Success and failure in learning to read: A critique of the research. Reading Research Quarterly, 1973, 8, 200-238.

Shankweiler, D., & Liberman, I.Y. Misreading: A search for causes. In J.F. Kavanaugh and I.G. Mattingly (Eds.), Language by ear and by eye. Cambridge, Mass: M.I.T. Press, 1972.

Smith, F. & Holmes, D.L. The independence of letter, word, and meaning identification in reading. Reading Research Quarterly, 1971, 6, 394-415.

Venezky, R.L. English Orthography: its graphic structure and its relation to sound. Reading Research Quarterly, 1967, 2, 75-106.

Wheeler, D. Processes in word recognition. Cognitive Psychology, 1970, 1. 59-85.

ACKNOWLE!)GEMENTS

The assistance of the Principal and staff of the Kennedy School and Our Lady of Victory School is acknowledge with gratitude

Support for this study was provided by the Joseph P. Kennedy, Jr. Foundation and the Spencer Foundation.

TABLE 1

ANALYSIS OF VARIANCE OF LATENCIES

Source	d.f.	Mean Square	F	d.f. for F	P
A (groups)	2	67.14	48.65	2,33	.001
subj. w. gps.	33	1.38	-----		
B (instructions)	1	229.02	233.6	1,33	.001
A x B	2	36.97	38.82	2,33	.001
B x subj. w.gps.	33	.98	-----		
C (quadrant)	3	.52	5.77	1,33	.05
A x C	6	.694	11.75	2,33	.01
C x subj. w. gps.	99	.091	-----		
B x C	3	.505	6.25	1,33	.01
A x B x C	6	.062	.75	2,33	n.s.
B x C subj.	99	.085	-----		

TABLE 2

POST-HOC COMPARISONS IN PHONEMIC COMPARISON

Comparison	Scheffe F	P
A1B2-A2B2	39.90	.001
A1B2-A3B2	135.3	.001
A2B2-A3B2	30.6	.001
B1-B2	32.02	.001
B2C1-B2C2	14.88	.01
B2C1-B2C3	25.0	.01
B2C2-B2C4	10.79	.05
B2C3-B2C4	19.61	.01

NOTE: A1-A2-A3: A1=Disabled; A2=Reading Matched
A3=Age Matched
B1-B2: B1=physical match; B2=sound match
C1-C2-C3-C4: Quadrants 1,2,3,4

TABLE 3

POST HOC-COMPARISONS IN CONTROL CONDITIONS

Comparison	Scheffe F	P
A1-A2	6.75	.05
A2-A3	20.56	.01
A1-A3	3.75	n.s.
B1-B3	13.97	.01
B2-B3	6.25	.05
B3 vs B1 and B2 at A2	14.33	.01
B3 vs B1 and B2 at A3	3.01	n.s.
B1 vs B3 at A3	3.52	n.s.

NOTE: A1=Disabled, A2=Reading matched, A3= Age matched
B1=Letters, B2=Non letter symbols, B3=word pairs

FIGURE CAPTIONS

Figure 1. Performance in the visual comparison test to letters in words.

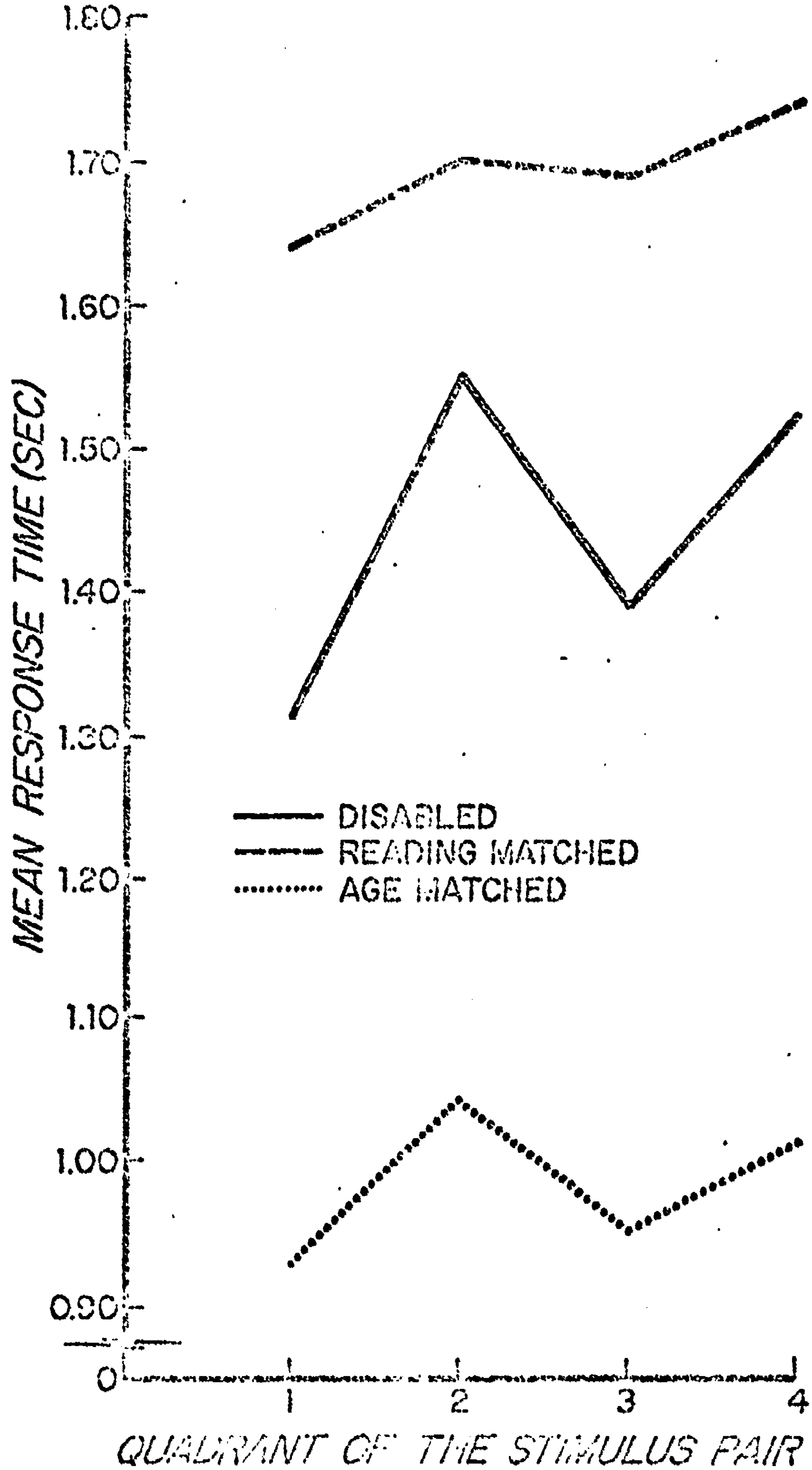
Figure 2. Performance in the vowel comparison task.

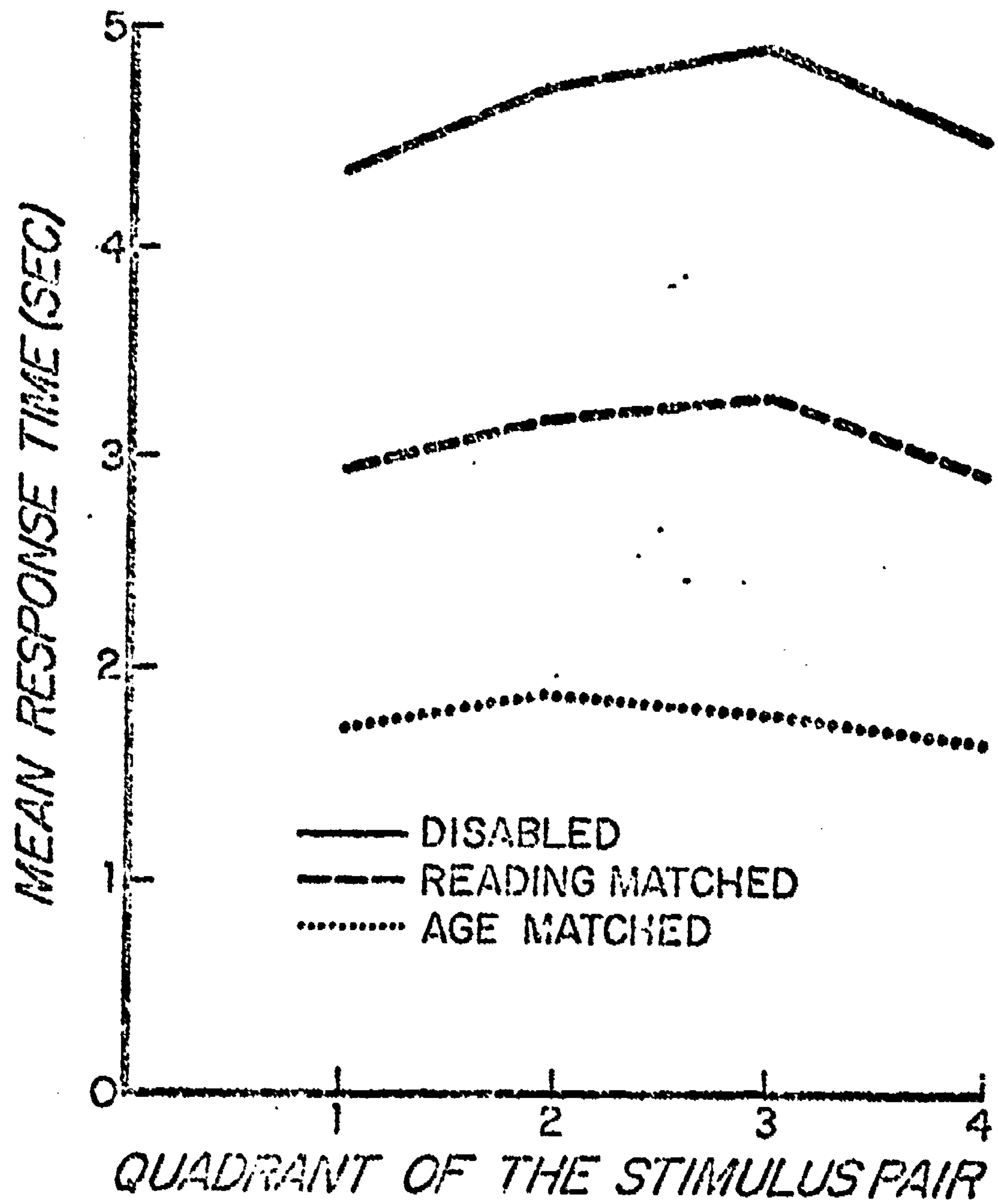
Figure 3. Performance in the visual control condition, in which latencies were averaged across "same" and "different" responses.

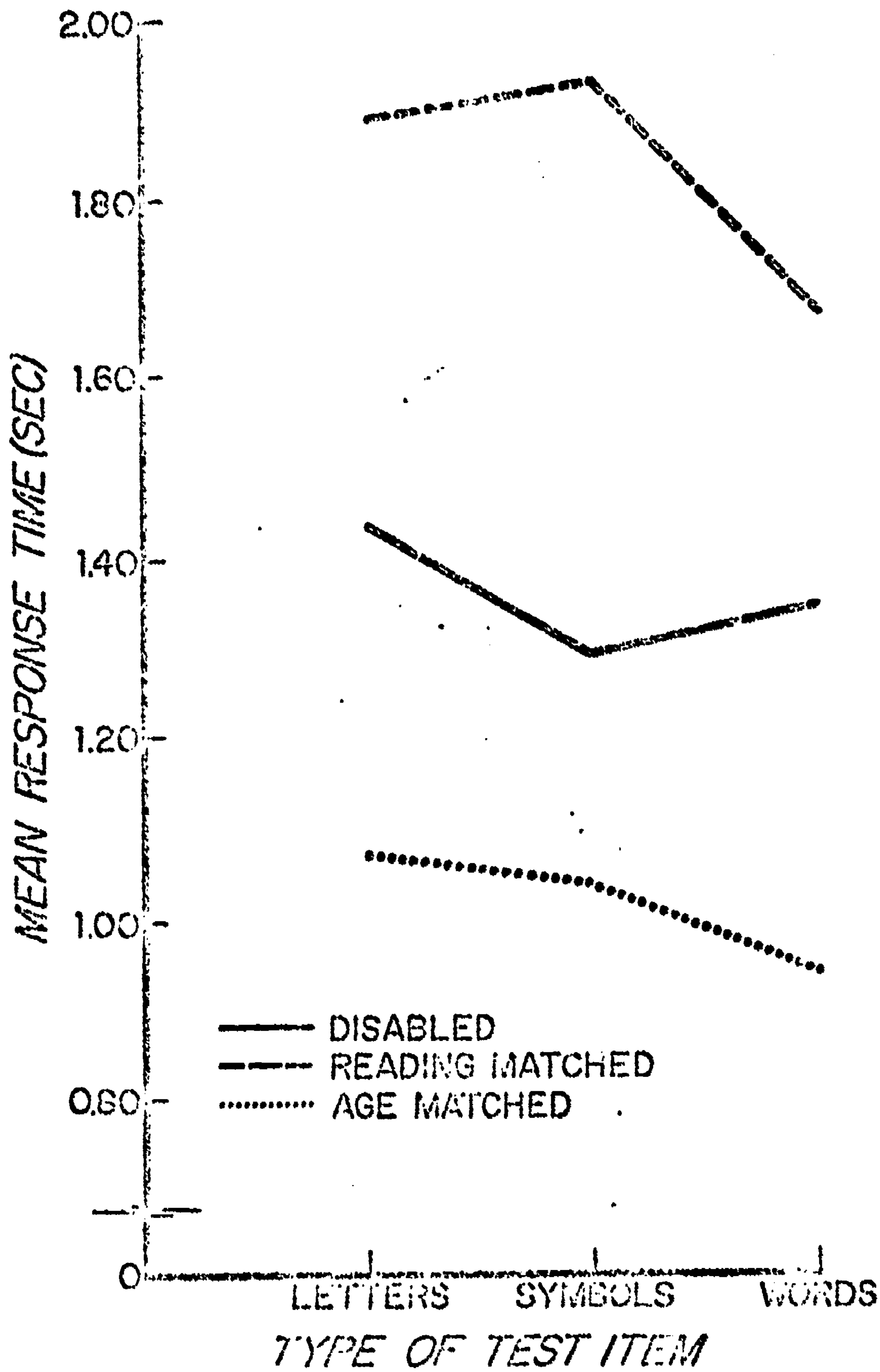
FOOTNOTE

1. Assistance in the conduct of this study was provided by the principal and staff of the Kennedy School for Special Learning Disabilities and the Principal and staff of Our Lady of Victory School.

Support for this experiment was provided by Maternal and Child Health Project Number 917, the Joseph P. Kennedy, Jr. Foundation, and the Spencer Foundation.


1.80
1.70
1.60
1.50
1.40
1.30
1.20
1.10
1.00
0.90
0
MEAN RESPONSE TIME (SEC)
DISABLED
READING MATCHED
AGE MATCHED
1
2
3
4
QUADRANT OF THE STIMULUS PAIR


MEAN RESPONSE TIME (SEC)
0
1
2
3
4
5
DISABLED
READING MATCHED
AGE MATCHED
1
2
3
4
QUADRANT OF THE STIMULUS PAIR



2.00
1.80
1.60
1.40
1.20
1.00
0.80
0
MEAN RESPONSE TIME (SEC)
DISABLED
READING MATCHED
AGE MATCHED
LETTERS
SYMBOLS
WORDS
TYPE OF TEST ITEM

READING COMPREHENSION DEFICIENCIES IN DISABLED READERS

Mary Seifert and John T. Guthrie

Johns Hopkins University

Abstract

Two studies were conducted each using a group of disabled readers and normal control groups. Sixty-two Ss were given IQ, sight vocabulary, oral reading and silent comprehension tasks. Both studies showed silent comprehension of disabled readers inferior to silent comprehension of controls matched on sight vocabulary and IQ ($p < .01$). Disabled readers were more variable than normals ($p < .01$) on an index of comprehension scores. Some disabled readers had normal comprehension skills while other Ss were deficient. Results indicated that comprehension deficiencies are identifiable in disabled readers and may be independent of other reading disabilities.

Published in the proceedings of the Congress of Learning Disabilities, 1973.

READING COMPREHENSION DEFICIENCIES IN DISABLED READERS

Reading is considered to be a developmental process, with mastery of some skills as a prerequisite to others. Gibson (1965) has postulated that "learning to differentiate graphic symbols and learning to decode letters to sounds" are skills acquired before the use of "progressively higher units of structure" (p. 1067). The skills needed for rendering letters into sounds and sounds into words are mastered before the complex process of grouping words and rendering them into meaning. The perception of characters, (the visual operations) precedes the development of a sensitivity to grammatical structure and the direct perception of the meanings of words in Kolers' Model (1970). Though these developmental stages suggest that word recognition skills are needed before comprehension skills are mastered, they do not suggest that comprehension is a natural outcome of correct word recognition.

Two methods have been used to measure reading comprehension. One of these is the question answer procedure. This method is common to standardized tests. It involves reading a passage; reading relevant questions and selecting appropriate answers. This means the Subject must first process the meaning of the passage and then he must process the meaning of the question. Finally, he must relate the question to the passage to choose an answer. The level of difficulty of the question is also a factor in this procedure. Different types of questions have different degrees of difficulty depending on their relationship to the passage (Bormuth, Manning, Carr and Pearson, 1970). A wrong answer could be the result of erroneous processing of the passage, of the question or of the relationship between them. It is evident that this procedure confounds the task of measuring reading comprehension.

The other method is the cloze procedure (Taylor, 1957) and its variations. This procedure involves the silent reading of a passage where words are randomly or selectively deleted and the subject's task is to fill in the blank with an appropriate word. Cloze has been used for linguistic analysis with normal reading children (Bormuth, 1966). In order to measure silent comprehension, the children must read silently with no feedback, filling in the blanks individually. This writing response requires a break in the normal, ongoing process of silent reading, and requires a shift of the emphasis from reading to the skill of writing. It also requires an ability to spell, which may be deficient in poor readers.

Two studies are presented which investigate silent reading comprehension, when correct word recognition is present, in good and poor readers. One purpose, also was to test a method which had been developed to ascertain the amount of comprehension occuring during silent reading.

Introduction - Study I

Good word recognition is often considered a prerequisite for reading comprehension. However, this does not necessarily mean that good word recognition skills will produce good comprehension skills or that poor word recognition skills will produce poor comprehension skills. To examine the effects of good and poor word recognition on comprehension, Oaken, Weiner and Cromer (1971) presented stories to good and poor fifth grade readers. The stories were presented under a good word identification condition in which good readers saw an unaltered reading passage and poor readers were taught the words in the passage. A poor word identification condition was included in which good readers saw a passage with spelling and grammar errors and poor readers were given no aid in word recognition. Results showed that the good readers' comprehension level was highest under the good word recognition condition. The good word recognition condition did not enhance the comprehension level of the poor readers over that of the poor identification condition. These results suggest poor readers do not extract meaning from reading passages in the same manner as normal readers. It is apparent that comprehension is not only word recognition, but also the classification of the words into meaningful syntactic relationships by the readers.

In studies of reading disability, disabled readers are usually matched with children their own age who are 1.5 years or more higher in reading level. It has been theorized by Satz, Rardin and Ross (1971) that older disabled readers also lag behind normal readers their own age in development of integrated language skills. They have found that young disabled readers (ages 7-8) do not differ in language abilities from young normal readers. Older normal readers (ages 11-12) displayed significantly higher language development than older disabled readers. The older disabled readers' language development, however, was not below that of the young normal readers. These similar levels of language development indicate that the reading comprehension performance of old disabled readers should be comparable to or higher than that of young normal readers.

The purpose of the following study was to look at silent reading comprehension abilities of good and poor readers at their respective reading levels.

Method

Subjects

There were three groups of subjects, 12 disabled readers, 12 normal old readers and 12 normal young readers. The disabled group had been screened to insure normal hearing, vision and emotional adjustment, absence of gross neurological dysfunction, a WISC IQ of 90 or above and 1.5 years or more deficiency in reading.

The normal old group was selected from a Catholic School and matched with the disabled group on chronological age and IQ, using the Peabody Picture Vocabulary Test. This group was also given the Gates-MacGinitie Reading Test, to insure that the group was also functioning at grade level. The normal young group was from the same school and matched the disabled group on IQ, using the PPVT, and reading level, using the Gates-MacGinitie Vocabulary Test.

There were no significant differences on IQ between the disabled group and the two normal groups. The young normals and disableds did not differ on reading level and the old normals and disabled did not differ on chronological age. All scores are presented in Table 1. Both groups of normal readers were judged by their teachers to display normal social, emotional, and academic adjustment.

Materials and Procedure

Seven passages were generated, based on stores from basal readers. The difficulties of the passages were primer - passage 1, 1^1-passage 2, 2^1- passage 3, 3^1-passage 4 based on Spache (1953) and 4^1- passage 5, 5^1- passage 6 and 6^1- passage 7 based on Dale-Chall (1948). The passage lengths were between 162 and 166 words. The passages were then written in maze form. In the maze form, a word selectively chosen from the story was displayed with two other words as:

The ground hog walked up the apple / near / hill .

The child read silently and circled the correct word. The maze task is a meaningful selection task within a context. The subject must look at the words surrounding the alternatives in terms of both meaning and correct grammatical structure.

The maze item always included: 1) the correct alternative (hill); 2) an incorrect alternative which violated the lexical requirements of the slot (near); 3) an incorrect alternative which violated the selectional requirements of the slot (apple). There were never fewer than 4 words between maze items, with 28 items per passage. The 2 incorrect choices for items were either on the Dale List of 769 Easy Words (1931) for passages primer through 3^1; or on the Dale-Chall List of 3,000 Familiar Words (1948), for passages 4^1 through 6^1. The 3 choices in each item were randomly ordered for all passages.

The disabled and young normal subjects were given the passages in groups of three, in three sessions, each of which never exceeded thirty minutes. The old normals were given the passages in groups of four, in two sessions, each of which never exceeded thirty minutes. None of the children had been exposed to the maze procedure before seeing the materials.

All Ss first worked through four examples with the examiner. They then completed the passages in order from primer to 6^1 with no feedback on accuracy. The same examiner administered all passages.

Results and Discussion

The comprehension differences between the three groups were analyzed. The correct responses that each S made on each of the 7 passages were analyzed with a 3 (groups) x 7 (passages) repeated measures analysis of variance. There was a significant effect for groups ($F = 63.45$, $df = 2/33$, $p < .01$), a significant effect for passages ($F - 4.09$, $df = 6/198$, $p < .01$) and a significant interaction effect ($F = 4.68$, $df = 12/198$, $p < .01$).

Post-tests with the Newman-Kuels procedure for differences between groups revealed that the old normals readers' performance was significantly higher than the young normals ($p < .01$) and the young normals were significantly higher than the disabled ($p < .01$).

The performance of each group on the passages was interesting. (See Figure 1). The normal old group performed at the 88 percent level or higher on all passages. The average reading vocabulary level for this group was 6.4. Comprehension performance on maze passage 7 (6^1 level) was 88 percent correct, ($\bar{X}$ = 24.64 items correct), which indicates that this group's comprehension level was comparable to their sight vocabulary level.

The normal young readers had a mean reading vocabulary level of 2.6. Looking at passage 3, at the 2^1 level, it can be seen that the average correct response was 64 percent. The disabled readers, with a mean reading vocabulary level of 2.5, had an average performance of 45 percent correct on passage 3 In fact, even on the simpliest passage, the primer passage 1, performance for both groups only reached the 70 percent level.

Thus the effect for passages showed that the old normal readers were the only Ss whose performance on the comprehension tasks was comparable to reading vocabulary level. This suggests that comprehension of sentences is a developmental process from the second to the fourth grade. The interaction effect indicates differences between the disabled and normal young readers, who were matched on reading level. Newman-Kuels post tests showed that the difference between performance on passage 1 was negligible. Also performance on passages 5, 6 and 7, which were beyond either groups' reading level, did not differ. Next, the three passages within the range of their reading vocabulary level, passages 2 (1^1), 3 (2^1) and 4 (3^1) were analyzed. The young normals performed significantly better than the disabled group ($p < .01$) on these passages. Thus though neither group had performance on the comprehension tasks comparable to their reading vocabulary level, the disabled group had quantitatively lower performance on these tasks. This suggests the presence of a deficiency in the development of comprehension skills for the disabled group. For the complete summary of results and further discussion see Guthrie (1973).

Introduction Study - II

The first study illustrated that disabled readers manifest reading comprehension deficiencies in comparison to both children who are their same age and children who possess their sight vocabulary level. The second study examined the effects of two variables on reading comprehension in normal and disabled readers: 1) amount of graphemic information and 2) oral reading ability.

Many contemporary models of reading contend that the reading process requires the reader to: 1) perceive the graphemes of words visually, 2) extract sounds from the graphemes and 3) derive semantic and syntactic cues from the sounds of the words, which enables the reader to "comprehend" word and sentences. (Kolers, 1970; Gibson, 1971, Goodman 1967). The focus of this study was on the relationships between stages 1 and 3 and stages 2 and 3. More specifically the questions posed for this investigation were a) how much graphemic information is necessary to extract meaning from a sentence and b) to what extent is word pronourciation indicative of the level of comprehension.

Methods

Subjects

The Ss were 13 normal second graders from a Catholic School, and 13 disabled readers from a non-public school. The disabled group had been screened to insure normal hearing, vision and emotional adjustment, absence of gross neurological dysfunction, a WISC IQ of 90 or above and 1.5 years deficiency in reading. The normal readers were judged by their teachers to display normal emotional, social and academic adjustment. The disabled group was matched with the normal readers on reading level and IQ. The Peabody Picture Vocabulary Test was used as an IQ measure with the normal and disabled groups having scores of 104.46 and 104.15 respectively. The Metropolitan Achievement Test Word Knowledge was used as a reading level measure with the grade equivalents being 2.73 for the normal readers and 2.51 for the disabled readers. The groups did not differ on these two criteria. Mean chronological ages of the normal and disabled groups were 7.60 and 9.61, respectively ($t = 6.28$, $df = 25$, $p < .0005$).

Materials and Procedures

The five silent reading passages were written in maze form, which was described in Study I. Words were selectively chosen from a basal style story. The story was then altered by adding two words to the chosen words displaying a maze item as follows:

The truck was full of on / roads. / corn

The child read silently and circled the correct word. In all tasks the alternatives included: 1) the correct alternative (corn); 2) an incorrect alternative which violated the lexical requirements of the slot (on); and 3) incorrect alternative which violated the selectional restrictions of the slot (roads). The three choices in the maze items in each of the passages were randomly ordered. The two incorrect choices were on the Dale List of 769 Easy Words (1931). The number of words in each of the five passages ranged from 162 to 165 and each passage was at the 2[1] grade level according to the Spache (1953) formula.

The five passages were developed in the following manner:

Passage 1: The maze items in this passage were made about every fifth word. There were 28 maze items in all. This passage was identical in form but not content to the passages in Study I.

Passage 2: This passage was the same in form but not content as passage 1 with the exception that the maze items were made approximately every tenth word, making it a 16 maze item passage.

Passage 3: There were 16 maze items in this passage as, in passage 2. Passage 3 was modified by deleting 12.5 percent of the letters of words, with the stipulation that the words modified or deleted (see passage 5) were not from maze items or from 2 successive words in the context. The procedure for the deletions was the omission of the last half of the letters of 25 percent of the words as:

```
                                  for
He put ears of co--   in   only lit--   truck.
                                  his
```

All deletions were noted by two dashes.

Passage 4: This passage also contained 16 maze items. The modification here was the deletion of 25 percent of the letters, following the stipulations noted for passage 3. In this case the last half of the letters of 50 percent of the words were deleted as:

```
                                   in
He go--   in hi--   truck and    went   do--  the  ro--  .
                                 picked
```

Passage 5: There were 16 maze items in this passage with 25 percent of the words omitted as:

```
                                  fast
The farmer and   ---   truck  --  rock  .
                                  much
```

In the second task the Ss read orally all of the words in random order, that were in Passage 1. The child received credit for pronouncing the word correctly. If he corrected an error before moving on to the next item, it was scored as correct.

The five passages were always given in the order of 1 to 5. The oral reading task was administered to half of the Ss prior to reading the passages and to half of the Ss after reading the passages. All tasks were administered by the same examiner.

Results and Discussion

Results were first analyzed to relate to the question of how much graphemic information is necessary for comprehension. The performance on passages 1 to 5 was compared. An analysis was conducted with percent correct on each task for each S as the dependent variable. A 2 (groups) x 5 (passages) repeated measures analysis of variance revealed a significant main effect for groups. ($F = 18.92$, $df = 1/24$, $p < .01$); a significant main effect for tasks ($F = 9.63$, $df = 4/96$, $p < .01$). Since there was no interaction effect, the total scores were analyzed across passages. Passage 1 will be discussed later.

Post tests with the Newman-Keuls procedure were performed. These showed that performance on passage 2, unaltered maze, was not significantly different from performance on passage 3, deletion of 12.5 percent of the letters. Performance on passage 4 (deletion of 25 percent of the letters) and 5 (deletion of 25 percent of the words) were significantly lower than performance on passage 2. Performance on passage 4 was inferior to passage 2 ($q = 6.0$, $df = 96$, $p < .01$) and passage 5 performance was inferior to passage 2 ($q = 6.91$, $df = 96$, $p < .01$). Deletion of 25 percent of the words or deletion of 25 percent of the letters of words clearly impairs comprehension of sentences. It appears that, however, 12.5 percent of the letters can be deleted without impairing sentence comprehension (Table 2). Young or disabled readers appear to need more than 75 percent of the letters present to adequately comprehend sentences.

The relationship between pronounciation level compared to silent comprehension of sentences was analyzed next. The oral reading task contained the words from passage 1. A ratio was constructed with performance on oral reading (percent correct) as the denominator. A ratio of less than 1.00 would suggest that the child could read orally material which he did not comprehend. A ratio greater than 1.00 would suggest that the child's comprehension surpassed his oral reading ability.

The normal group had a mean ratio of .99 and a standard deviation of .09 with only 1 normal reader having a ratio lower than .90. (see Figure 2). This result suggests that for normal readers sentence comprehension is very closely aligned with their ability to orally read the words in the sentences. Also, since the mean performance for this group on maze and oral reading was 89.23 and 89.51 respectively, (Table 2) a ceiling effect did not influence these findings.

The ratio for disabled readers gave a mean of .72 and a standard deviation of .21. This mean was significantly lower than the normal mean ($t = 4.50$, $df = 25$, $p < .01$). An inspection of Figure 2 shows that this group was divided into thirds at points: .50, .70, and 1.00. As a group the disabled readers did not comprehend the sentences as well as normals. When examined individually, it is clear that for one third of the disabled readers, comprehension level was comparable to oral reading ability. However, two-thirds of the disabled readers were reading materials which they did not comprehend.

These findings suggest that oral reading performance is quite likely to be a sufficient indication of silent reading comprehension performance for normal readers, but not for disabled readers.

Summary

Two studies have examined reading comprehension abilities in good and poor readers. The first study used materials of increasing difficulty. The results of this study suggested that comprehension is a developmental process. Old normal readers performed significantly higher than young normals and disabled on all passages, including the passages within the reading range of these 2 groups. Also comprehension level was equal to reading vocabulary level only for the old normal readers. Though the young normal readers did not perform the comprehension tasks at their reading level, their performance was significantly higher than the disabled readers, who had the same reading level. This suggests the presence of a reading comprehension deficiency in the disabled group as opposed to a developmental lag in these skills.

The relationship between comprehension and the amount of graphemic information needed and used by good and poor readers was investigated in the second study. Both groups needed more than 75 percent of the letters of the words to extract meaning from sentences in which the words were located, although they could comprehend the sentences with 12.5 percent of the letters omitted. This indicates that at the second grade reading level, a substantial amount of graphemic information is necessary for comprehension. From this data, it cannot be determined whether the Ss decoded the partial words into sounds and derived meaning from the sound or whether they derived meaning directly from the partial words.

Normal readers comprehended sentences to the same extent that they could read them orally. If a child could read 75 percent of the words, his maze performance was at about the 75 percent level. However, disabled readers did not comprehend sentences to the same extent that they could read them orally. Level of performance for two-thirds of this group was higher on oral reading than on comprehension. This group used the graphemic information available to pronounce words but not to extract meaning.

Both studies showed the disabled readers to have quantitative deficiencies in comprehension. They also both indicated that word recognition is not sufficient for sentence comprehension for most disabled readers. Normal readers process material that they can phonetically decode in such a way to produce meaning.. Disabled readers appear to be unable to do this. There is a need for instruction designed to ameliorate comprehension deficiencies in addition to instruction in visual perception and phonetic decoding.

REFERENCES

Bormuth, J.R., Readibility; a new approach, Reading Research Quarterly, 1966, 1, 79-132.

Bormuth, J.R., Manning, J., Carr, J., Pearson, D., Children's comprehension of between and within sentence syntactic structures. Journal of Educational Psychology, 1970, 61 (5), 349-357.

Dale, E., A comparison of two word lists. Educational Research Bulletin, 1931, 10, 484-489.

Dale, E., and Chall, J., A formula for predicting Redability. Educational Research Bulletin, 1948, 27, 11-20; 37-54.

Gibson, E., Learning to read. Science, 1965, 1066-1072.

Gibson, E.S., Perceptual learning and the theory of word perception. Cognitive Psychology, 1971, 2, 351-368.

Goodman, K.S., Reading: a psycholinguistic guessing game. Journal of the Reading Specialist, 1967, 6, 126-133.

Guthrie, J., Reading comprehension and syntactic responses in good and poor readers. Journal of Educational Psychology, in press.

Kohlers, P., Three stages of reading in J.R. Williams and H. Levin, (Eds.). Basic Studies on Reading, New York: Basic Books, Inc., 1970.

Oaken, R., Weiner, M., Cromer, W., Identification, organization and reading comprehension. Journal of Educational Psychology, 1971, 62, 71-78.

Satz, P., Rardin, D., Ross, J., An evaluation of a theory of specific developmental dyslexia. Child Development, 1971, 42, 2009-2021.

Spache, G., A new readability formula for primary grade reading materials. The Elementary School Journal, 1953, 53, 410-413.

Taylor, W.L., "Cloze" readability scores as indices of individual differences in comprehension and aptitude. Journal of Applied Psychology, 1957, 41, 19-26.

TABLE 1

CHARACTERISTICS OF SUBJECTS IN STUDY I

Subjects	Chronological Age		IQ[1]		Reading[2] Level	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Normal Old	10.12	.45	105.58	8.74	6.43	1.68
Normal Young	7.42	.46	108.33	12.00	2.62	.85
Disabled	9.99	1.36	100.08	10.85	2.54	.40

1. Peabody Picture Vocabulary Test IQ
2. Gates-MacGinitie Vocabulary Test, Grade Level

TABLE 2

PERFORMANCE ON ORAL READING AND

MAZE TASKS IN STUDY II

SUBJECTS		$\frac{\text{MAZE}}{\text{ORAL}}$	ORAL READING PASSAGE I	PASSAGES				
				1	2	3	4	5
Normal Readers N = 13	$\overline{X}$	.99*	89.54*	89.23*	84.38*	81.38*	71.85*	67.85*
	SD	.09*	7.11	11.81	12.26	17.99	19.99	24.88
Disabled Readers N = 13	$\overline{X}$	.72	80.92	60.62	53.10	47.85	45.85	46.85
	SD	.21	12.91	24.88	25.57	24.39	18.89	16.54
Total	$\overline{X}$	.72		74.93/	68.74	64.62	58.85/	57.35/
	SD			24.02	23.32	27.08	22.28	20.92

NOTE: Scores in the $\frac{\text{Maze}}{\text{Oral}}$ column are ratios. Other scores are percent correct.

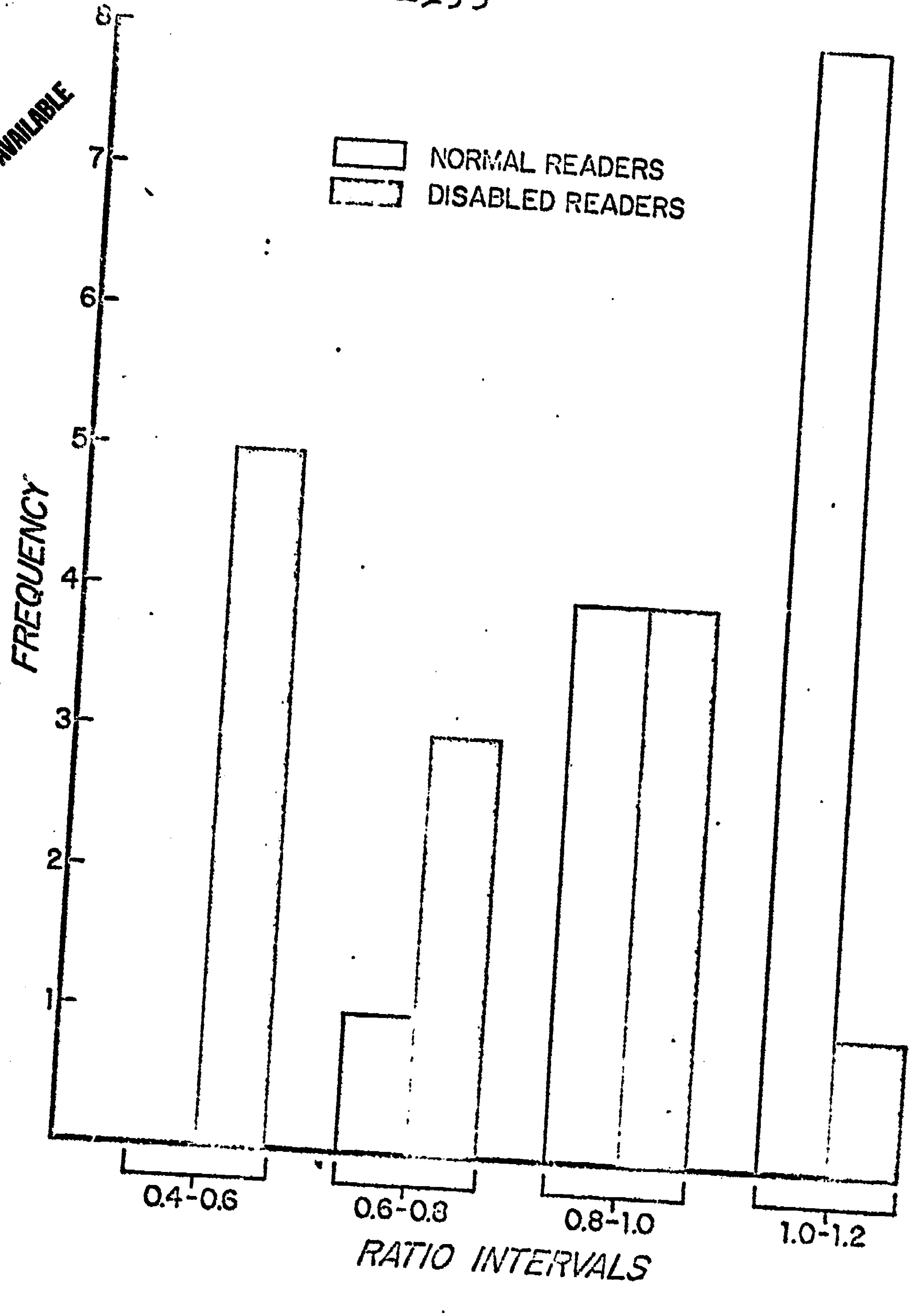
* Normal and Disabled groups differed at $p < .01$.

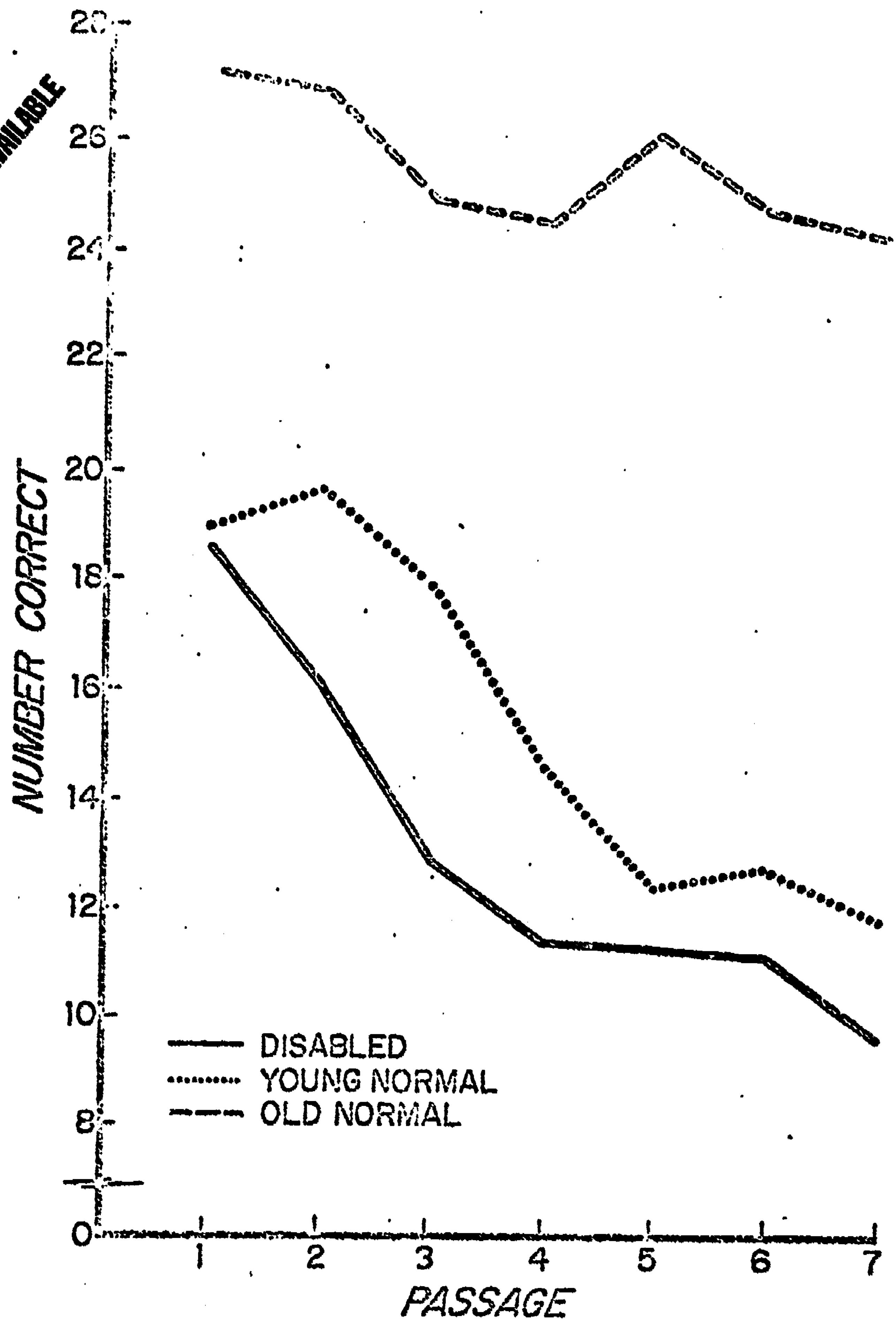
/ Indicates that the task differs from task 2 a $p < .01$.

FIGURE CAPTIONS

Figure 1 Reading comprehension as a function of passage difficulty and group in Study I.

Figure 2 Distribution of maze/oral reading ratios in normal and disabled readers in Study II


NORMAL READERS
DISABLED READERS
FREQUENCY
8
7
6
5
4
3
2
1
0.4-0.6
0.6-0.8
0.8-1.0
1.0-1.2
RATIO INTERVALS


NUMBER CORRECT
28
26
24
22
20
18
16
14
12
10
8
0
1
2
3
4
5
6
7
PASSAGE
DISABLED
YOUNG NORMAL
OLD NORMAL

PERCEPTION AND MEMORY IN THE PERFORMANCE OF NORMAL AND DISABLED READERS ON AN AUDITORY DISCRIMINATORY TASK[1]

Timothy Sheehan[3] and John T. Guthrie[2]

THE JOHNS HOPKINS UNIVERSITY

ABSTRACT

The performance of 16 normal readers and 16 disabled readers on an auditory task was studied in an attempt to expose the components of perception and memory underlying such tasks. A task was developed in which items conformed to a delayed match to sample paradigm. A standard was followed by three alternatives; words taken from the Wepman Test of Auditory Discrimination were used as stimuli. The distribution of errors among alternative positions was used to distinguish the components of perception and memory. In an 2 X 2 X 3 analysis of variance, the main effects of reading ability and prior exposure to the standard words, the experimental variable, were significant at (p .01). Prior exposure interacted with position of alternative in the group of alternatives (p .025). It was concluded that the errors on the task are attributable to the perceptual rather than the memory component of the task; and, secondly, that the deficit exhibited by disabled readers is almost exclusively a perceptual deficit.

72-124

PERCEPTION AND MEMORY IN THE PERFORMANCE OF NORMAL AND DISABLED READERS ON AN AUDITORY DISCRIMINATORY TASK

Timothy Sheehan and John T. Guthrie

One approach to the investigation of reading disability has been to assess the performance of normal readers and disabled readers on measures of the particular information processing capacities thought to underlie the reading process or its acquisition. Comparison of samples of NRs and DRs is expected to reveal specific areas of deficiency among disabled readers. Auditory discrimination is a processing capacity that has been investigated in this manner and shown to be an area of deficiency for DRs (Wepman, 1960; Deutsch, 1964; Blank, 1968).

The Wepman Test of Auditory Discrimination (ADT) was the instrument employed in each of these investigations. This measure consists of 40 pairs of words, 10 in which the words are identical and 30 in which the words differ in a single phoneme. The subject renders a "same-different" judgment following the audition of each pair. This task, simple enough on the surface, involves a complex procedure on the part of the subject; he must perceive each word, hold it in memory, and make the necessary comparison. Breakdown occuring at any phase of this operation may result in error. Therefore, auditory discrimination as assessed by the Wepman ADT is no simple capacity at all but a complex procedure involving the integration of several capacities.

The present study undertook to examine the perceptual and memory components of the auditory discrimination process and to gauge the contributions made by the separate components in the performance of normal and disabled readers. This objective required that the task be modified, yet not so extensively that generalization of results to the Wepman ADT would not be warranted.

Since previous studies have reported that disabled readers do not exhibit a deficit when required to identify single word items (Doehring and Rabinovitch, 1969; Blank, 1968), the present task included more than one stimulus per item. The stimuli were words drawn primarily from the Wepman ADT. The needs of the present study required that items contain more than the single discrimination demanded by the items of the Wepman ADT. Items of the present task assumed a temporal structure analogous to the spatial pattern of this example:

SHOP SHOT SHOP SHOCK

S was presented a standard and three alternatives in succession. S was required to indicate each instance of a standard to alternative match immediately after it occurred.. Any number of matches could occur in a single item.

Perceptual and memory components are differentiated by examining the distribution of errors over positions. In this interpretation errors based on the inaccurate perception of a stimulus distribute evenly over positions. This is because erroneous judgments that reflect the congruence or dissonance of an alternative with respect to a faulty perception of the standard depend upon the peculiar faults of the perception and not upon the position of the alternative. Similarly, there is no reason to suppose that an alternative is more likely to be inaccurately perceived as a result of the particular position it occupies.

A memory trace deteriorates either by decay over time or by interference from interpolated activity. Thus, errors attributable to forgetting increase over positions. Two points will be noted concerning this analysis. First, forgetting affects only the standard: memory for the alternative prior to comparison is not distinguished from the perception of the alternative. Second, errors arising from inadequate comparison of alternative to standard and errors arising from inaccurate perception of an alternative are confounded.

This interpretation can be phrased in terms of a linear function describing the distribution of errors over positions. The contribution of the perceptual component is estimated from the intercept; the memory component from the degree of positive slope. Figure 1 provides a graphic illustration of the linear model. In the present analysis the frequency of first position errors is treated as the intercept, while evidence concerning the degree of positive slope is found in the comparison of the frequency of errors in the first and the third positions.

The linear model assumes that the combination of the perceptual and memory components proceeds in an additive manner. Since the contribution of the perceptual component is constant, the intercept estimates the strength of that contribution at any position. Therefore, if errors are more frequent at the second and third positions than at the intercept, forgetting is inferred. Any such increase in error frequency will be directly reflected in the slope of the linear function.

Prior exposure to the standard words was selected as a treatment variable because there was good reason to believe that prior exposure would improve performance by facilitating the perception of the standard. Of particular interest would be differences between normal and disabled readers either in the size of the gains realized in the treatment condition or in the manner, in terms of the perceptual and memory components, in which the gains were realized.

Method

Materials and Apparatus

Items were constructed to maintain stimulus similarity with the Wepman ADT. From each of the pairs on Form II of the Wepman ADT one word served as the standard the other as the alternative in the same item. Since additional alternatives were needed, the principal of "minimal pairs" upon which the Wepman ADT is constructed was applied. By this principle "different" words differ from the standard in only one phoneme. This difference is further limited to discrepancies within a minimal number of phonetic (articulatory) categories.

The words were recorded on a tape recorder. The original tape was then re-recorded with volume modulation to equate stimuli within items on peak volume. Measured volume differences among stimuli within items were held to less than 2 db. in all but two instances. Thus, the effectiveness of volume as a cue was restricted severely.. The Ss wore headphones and responded by depressing a button which ignited a light to indicate a standard to alternative match. The button presses were recorded E.

Subjects and Design

The normal sample was drawn from a parochial school in Baltimore; the disabled sample came from a population of disabled readers who were attending a remediation program at the Kennedy Institute. Normal readers were administered the comprehension and word knowledge subtests of the Metropolitan Achievement Test Primary II and the Peabody Picture Vocabulary Test which gives an IQ score. Hearing was screened on a standard audiometric test given on a Beltone audiometer.

The population of disabled readers were defined by these criteria: reading performance of 1.5 years below expectation computed from CA; completion of first grade; lack of gross neurological, sensory, or physical handicaps; and IQ greater than 90 as measured by the WISC. The samples were matched for age and IQ; the figures are presented in Table 1. The differences on IQ and on age were not significant.

Two between S variables, reading level and prior exposure (PE), were combined with one within S variable, position of alternative in a 2 X 2 X 3 factorial design with repeated measures. Ss were ranked by reading level within the reading groups, pairs were formed between adjacent Ss on this scale, and one member of each pair was assigned at random to the experimental conditions PE and no prior exposure (NPE).

<u>Procedure</u>

A single session was sufficient for all Ss. PE consisted in the presentation of the 40 standard words. In addition a picture of each word on an index card was presented. A set of five words and corresponding pictures was presented. The S responded by repeating the word and indicating the proper pictorial associate. When the S reached a criterion of correct performance on all items in succession in one set, the next set was presented.

Ss in condition PE were informed that the words to which they had been exposed would appear as standards in the test. Otherwise all Ss were similarly instructed on the auditory match to sample task. The task was described: words would appear in groups of four, listen carefully to the first word, press the button each time this word is repeated in the next three words, then get ready for the next group. E assisted S in performing the first example. S then performed alone a second example. If he failed, instructions were elaborated until two examples were successfully completed.

Results

Two types of errors may occur in a match to sample task. A match may be indicated when in fact no match exists; or, a match may, in fact, exist and fail to be indicated. These two types of errors, false positives and false negatives, respectively, were combined to yield a single dependent variable. An analysis of variance was performed for a 2 X 2 X 3 design with the third factor, position of alternative, treated as a repeated measure. Criterion levels of significance were attained by the main effects for Reading ability (F=21.907, df=1, 28, p .001); and for PE (F=8.057, df=1, 28, p .01). One interaction attained significance, PE X Position (F=4.521, df=2, 56, p .025). The means and standard deviations for these data are presented in Table II.

The direction of the main effects was that normal readers were superior to disabled readers and Ss with prior exposure made fewer errors than Ss without prior exposure. The nature of the PE X Position interaction was that prior exposure improves performance in the first and second alternative positions but not in the third. This can be observed in Figure 2. This interpretation was verified by the use of a post hoc test for simple main effects of PE at each alternative position (Winer, 1962). The effect of PE was found to be significant at positions one and two (F=51.5, df=1, 28, p .001; and F=14.15, df=1, 28, p .05, respectively); but not at position three (F 1). In addition a Newman-Keuls post hoc test (Winer, 1962) revealed that in condition PE significantly more errors were made at position three than at position one; while the reverse held true for condition NPE, more errors were made at position one than at position three (q=2.83, fd=2, 56, p 05; and q=3.19, df=2, 56 p .05, respectively).

Discussion

These results form a sensible pattern when they are arranged according to the linear model. Recall that in the slope, intercept analysis the perceptual component is estimated from the frequency of first position errors, while the memory component is revealed in an increase in errors from the first to the third positions. Viewed in these terms the data yield two noteworthy findings: first, the perceptual component was the predominant source of errors; second, this held true for both normal and disabled reading groups. It follows from these findings that the deficit exhibited by disabled readers on the task is accountable in terms of perception.

The salient aspect of the data is the lack of compelling evidence that forgetting was a significant source of errors for either reading group. The lack of significance of the main effect of position and the interaction of position with reading ability in the analysis of variance indicates that there was little change in error frequency over positions in the performance of either reading group. In the absence of evidence for a rise in errors over positions forgetting is discounted as a significant source of errors. Differences in performance, then, between samples can be adequately explained in terms of perception. Expectations regarding the effect of prior exposure to the standard words were confirmed. PE resulted in significant gains over NPE. Evidence that these gains were mediated by the facilitation of perception is found in Figure 2. If the effect of PE had been the strengthening of the memory trace, there would have been a reduction of third position errors relative to first position errors. This did not occur. Third position errors were unchanged by PE, while first position errors were reduced dramatically. In effect, a relative increase in third position errors occurred. Perceptual facilitation indicated by the significant reduction of errors in the first position emerges as the sole mediator of PE effects.

The failure of PE to interact with reading ability indicates the generalizability of PE effects across reading groups. That is, there is no evidence that prior exposure affected normal and disabled readers differently either in the size of the gains realized or in the manner in which these gains were achieved.

The deficit exhibited by retarded readers on the present task was not unexpected in view of previous research on tasks of auditory discrimination. It seems reasonable to assume that the results obtained with the Wepman ADT are qualitatively similar to those obtained with the present instrument. High stimulus similarity was insured by making direct use of the Wepman ADT materials. Response similarity is less straightforward. However, delayed match to sample items, such as those used here require a succession of private "same-different" judgments. Thus, an equivalence exists between the implicit response on the present task and the explicit response on the Wepman ADT. The present findings may therefore be generalized to the Wepman ADT.

The significant decrease in errors from the first to the third positions in the NPE condition poses a problem for the additive linear model. A decrease in errors over positions is explainable only if the notion that the perception of the standard improves over time is accepted. The present study was not designed to test that notion and provides no basis for its evaluation. In addition, there is the indication of significant forgetting in the PE condition. This can be resolved in one of two ways: either forgetting is unique to that condition, that is, is somehow engendered under that condition; or, forgetting is ubiquitous, but is only discovered when perceptual errors are reduced by PE. Of these the latter is more plausible. The acceptance of either, however, would not mitigate the conclusion that the predominant source of errors is the perceptual component of the task.

REFERENCES

Blank, M. Cognitive processes in auditory discrimination in normal and retarded readers. Child Development, 1968, 39, 1091-1101.

Deutsch, C. P. Auditory discrimination and learning: social factors. Merrill-Palmer Quarterly, 1964, 10, 277-296.

Doehring, D. J. and Rabinovitch, M.S. Auditory abilities of children with learning problems. Journal of Learning Disabilities, 1969, 2, 467-475.

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

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

FOOTNOTES

1. Partial support for the project was from The Joseph P. Kennedy, Jr. Foundation.

2. The cooperation and assistance of Sister Marie, Principal, Sister Carolyn and Mrs. Martinez, teachers and Mrs. Geraghty, secretary of the Our Lady of Victory School in Baltimore are gratefully acknowledged. Appreciation is also expressed to Malcolm Preston and Rachel Stark for their advice concerning the preparation of the test materials.

3. Requests for reprints should be addressed to Timothy Sheehan, Department of Education, The Johns Hopkins University, Baltimore, Maryland 21218.

TABLE I

STATISTICAL SUMMARY OF SAMPLES

		IQ	CA	MAT Comprehension
Normal Readers (n=16)	$\bar{X}$	113.0	9.08	4.49
	SD	8.52	.41	.37
Disabled Readers (n=16)	$\bar{X}$	106.75	9.18	1.93
	SD	12.77	.99	.40

TABLE II

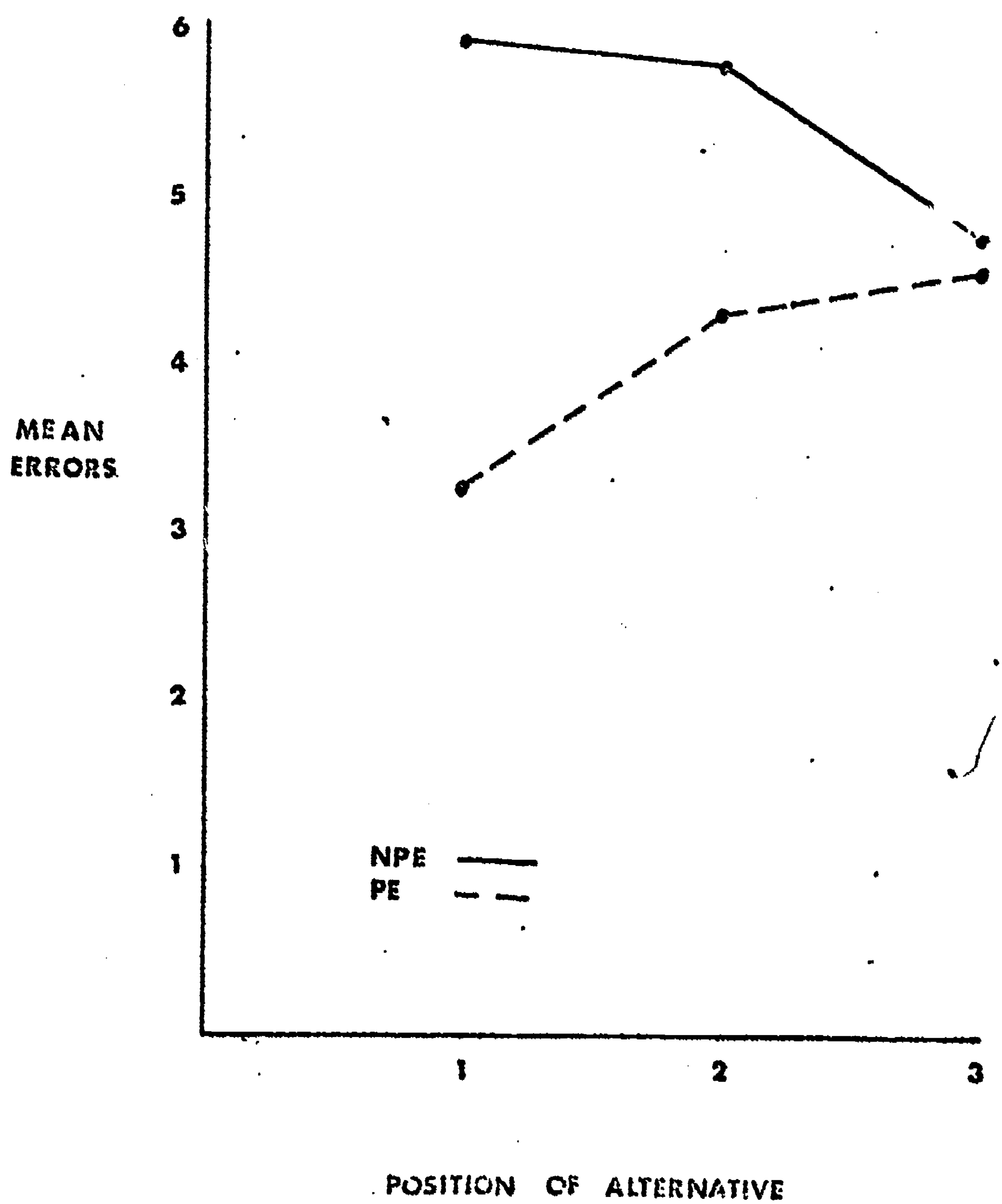
CELL MEANS AND STANDARD DEVIATIONS

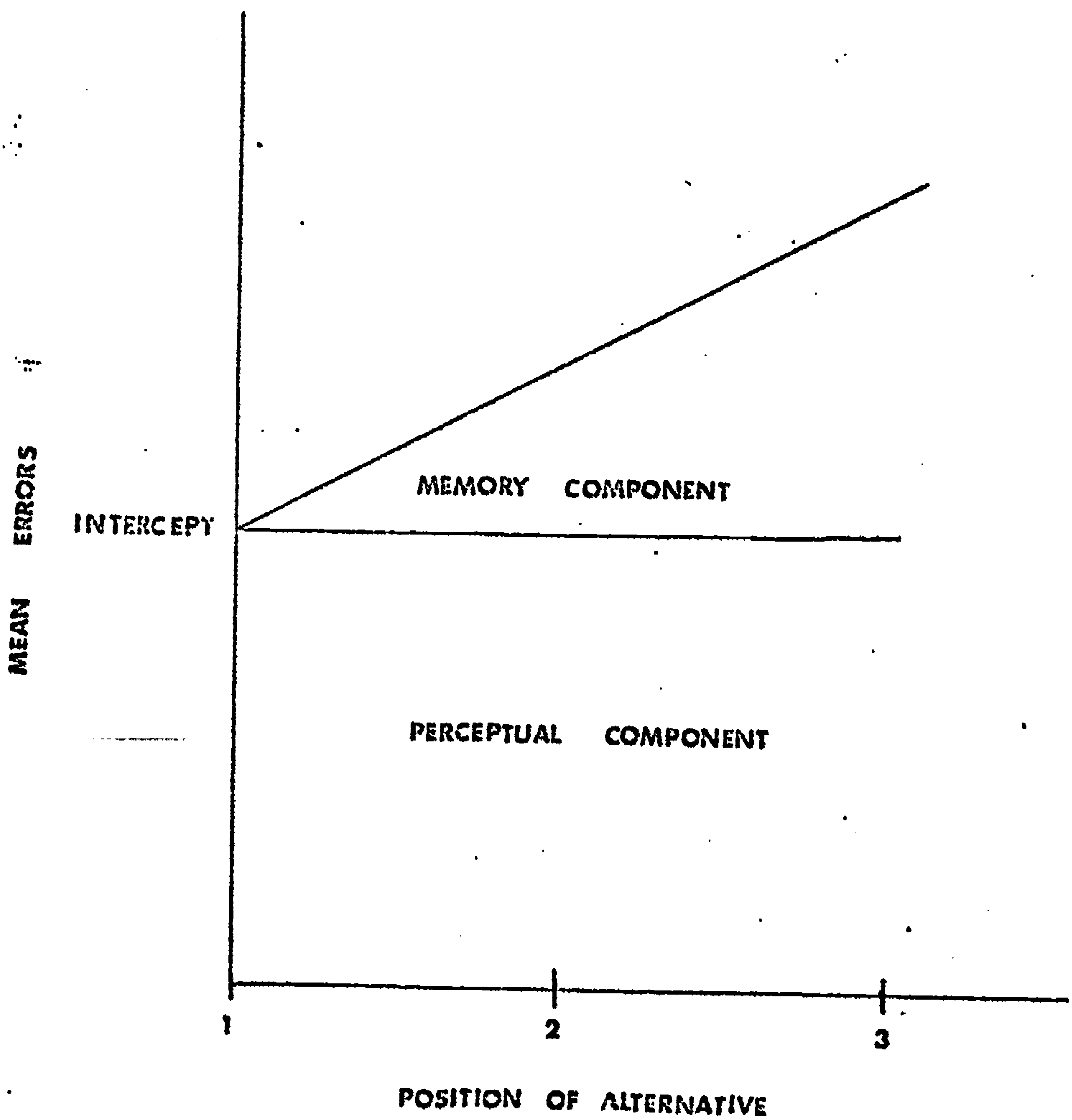
		1	2	3
Normal Readers	PE $\bar{X}$	2.43	2.71	3.86
	SD	.98	1.98	1.93
	NPE $\bar{X}$	4.78	3.89	3.78
	SD	1.79	.93	1.48
Disabled Readers	PE $\bar{X}$	4.25	5.63	5 13
	SD	2.92	1.69	.83
	NPE $\bar{X}$	7.35	8.00	5.88
	SD	3.06	2.20	2.75

FIGURE CAPTION

Figure 1 Additive linear model of performance on the auditory discrimination task.

Figure 2 Interaction of prior exposure and position.


6
5
4
3
2
1
MEAN
ERRORS
NPE
PE
1
2
3
POSITION OF ALTERNATIVE


MEAN ERRORS
INTERCEPT
MEMORY COMPONENT
PERCEPTUAL COMPONENT
1
2
3
POSITION OF ALTERNATIVE