DOCUMENT RESUME

ED 046 627 | 24 | RE 003 195

AUTHOR	Calfee, Robert C.; And Others
TITLE	How a Child Needs to Think to Learn to Read.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	TR-131
PUB DATE	Jul 70
CONTRACT	OEC-5-10-154
NOTE	48p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Auditory Tests, *Educational Testing, Kindergarten Children, Listening Skills, *Prereading Experience, Reading Readiness, *Reading Readiness Tests, *Reading Skills, Standardized Tests, Visual Measures, Vocabulary Skills

ABSTRACT

The component prereading skills of kindergarten children in two Wisconsin cities were examined. It was assumed that independent component skills exist and that investigation of separate skill areas would point out relevant combinations. Selected for study were visual, acoustic-phonetic, letter-sound association, and vocabulary skills. The research strategy included the compilation of a basic skills tests battery and individual administration of the battery to 21 middle-class and 22 lower-class subjects. The test results are reported separately for each skill area and for correlated skills. Predicted relationships were found to exist between visual matching and alphabet knowledge and between the acoustic-phonetic tasks of segmentation and rhyme-production. Other unpredicted correlations were found, but in general the tests appeared to be independent of one another. A bibliography and tables are included. (MS)

HOW A CHILD NEEDS TO THINK TO LEARN TO READ

WISCONSIN RESEARCH AND DEVELOPMENT

CENTER FOR COGNITIVE LEARNING

0001

Technical Report No. 131

HOW A CHILD NEEDS TO THINK TO LEARN TO READ

by
Robert C. Calfee, Robin S. Chapman, and Richard L. Venezky

Report from the Project on Basic
Pre-Reading Skills: Identification and Improvement

Robert C. Calfee and Richard L. Venezky
Principal Investigators

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

July 1970

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Project on Basic Pre-Reading Skills: Identification and Improvement in Program 1. General objectives of the Program are to generate new knowledge about concept learning and cognitive skills, to synthesize existing knowledge, and to develop educational materials suggested by the prior activities. Contributing to these Program objectives, this project's basic goal is to determine the processes by which children aged four to seven learn to read and to identify the specific reasons why many children fail to acquire this ability. Later studies will be conducted to find experimental techniques and tests for optimizing the acquisition of skills needed for learning to read.

ACKNOWLEDGMENTS

Special thanks are due to the experimenters who carried out the actual testing: Ella Alexander, Miriam Gelman, Stephen Harrington, Claudia Shellenberg, Michael Vickers, and Nancy Wellman.

CONTENTS

LIST OF TABLES AND FIGURES

ABSTRACT

This report summarizes 1968-69 studies of component pre-reading skills of kindergartners. Working from the assumption that independent component skills might exist, four skill areas were selected for investigation on the basis of literature findings and a decoding model of reading: visual, acoustic-phonetic, letter-sound association, and vocabulary skills. A research strategy is presented for the development and validation of assessment tests in each area and the subsequent development and validation of training procedures. The research reported here is concerned with skill assessment in each of the areas.

The four skill areas were investigated through the use of experimental Basic Skill Test batteries administered individually to kindergartners from a lower-middle class population in Madison (N = 21) and a lower class population in Beloit (N = 22). Madison Ss participated in three 30- to 50-minute sessions in November, December, and January; Beloit Ss in one 45-minute session in March.

Methods, results, and discussion are presented in detail for each of the four skill areas and invalid test procedures identified (e.g., same-different tests of rhyming). In the remaining tests, errors are examined for sources of difficulty (e.g., left-right letter reversals) and the pattern of intertest correlations examined. The findings support an assumption of independent skills related to reading.

I
INTRODUCTION

For the past 3 years, the authors have been engaged in the study of cognitive skills related to the acquisition of reading. The primary goal of the research has been increased knowledge about the early stages in the development of the reading process, focusing on those aspects of early reading which involve decoding or translating from the written to the spoken form of the language, as opposed to what is commonly referred to as "comprehension."

The plural, cognitive skills, is appropriate for two reasons. First, reading is not a unitary ability but a collection of abilities which develop rapidly during the primary grades. For example, the first grade child, slowly translating strings of abstract visual symbols into the familiar spoken language, is almost certainly operating in a different manner than a college student scanning through a text. Second, each of these abilities requires of the child a variety of prerequisite cognitive skills. As will become obvious, from our point of view cognitive skills are very specifically defined: recognizing that two visual forms are identical, locating a test form in an array, or detecting whether two spoken words rhyme. For other investigators, cognition is virtually synonomous with theory or knowledge; our thinking about cognition has been closer to the approach which has come to be called "human information processing."

A second aim of our research project is to use knowledge about the reading process gained through basic research on component cognitive skills to direct the development of more effective procedures for teaching reading than are now available. It should be stressed that we are not looking for a new and better way of teaching, *per se*. The substantial body of research on the relative efficiency of various methods of reading instruction has not produced many exciting results (e.g., Chall, 1967; Bond & Dykstra, 1967). It still takes too long for most children to achieve a reasonable level of competence in reading, and there are still far too many school children who are functionally illiterate. To repeat the point made by Eleanor Gibson (1965), "good pedagogy is based on a deep understanding of the discipline to be taught and the nature of the learning process involved (p. 1072)."

This paper will not concern itself with the second aim above of improving reading instruction, because we have yet to grasp that "deep understanding." For those who like to speculate about how reading ought to be taught, there is adequate stimulation in the technical and popular literature. Instead, this Technical Report will focus on the methodological base from which we are currently working.

Attacking a problem like "how does a child learn to read?" is an interesting challenge to those of us with a bias toward reductionism. Even when the problem is limited to the decoding stage, it is obvious that there are a multitude of task demands—visual and auditory perception, attention, learning, inductive reasoning, and so on—any of which might easily absorb one for a lifetime of investigation. How does a person carry out research that is at both extensive and intensive? There are compelling social reasons for trying to locate the most significant "hang-ups" in reading instruction before another generation passes. Yet it is all too easy to deal superficially with these problems. Scanning the literature on reading, one becomes convinced of two things: (a) almost any aspect of a child's intellectual performance correlates to some extent with reading achievement and (b) almost any remedial effort will be of some help to the child in learning to read. In fact, the most relevant variable in determining the success of any remedial program would appear to be time—the amount of time a teacher spends in one-to-one contact with an individual child (Monroe, 1932, p. 150).

Our research decisions have been guided by certain decisions about the innovative potential of various areas. First, we have decided to concentrate on the initial phases of reading instruction, including reading readiness. Accordingly, we are most interested in the performance of children between 4 and 7 years of age. Second, we have decided that those children with the poorest prognosis represent the greatest opportunity for improvement, although they also pose the greatest challenges in research.

To those readers who are experimental psychologists, these introductory remarks may strike a strange note. Not since Tinker's (1958) paper on eye-movements in reading has there been an article in either the Psychological Bulletin or Psychological Review concerned with reading. Just as B. F. Skinner bemoaned the "flight from the laboratory," so there might be cause for concern about the absence of learning psychologists from the classroom. This situation has begun to change lately; for example, reading research programs have been established in the last few years under the supervision of "hardheaded" experimentalists such as Eleanor Gibson, Harry Levin, Richard Atkinson, Harry Silberman, and Robert Glaser, to mention a few. In addition, current research in human information processing, psycholinguistics, and developmental cognitive psychology has had considerable influence on our work on reading as will be pointed out in connection with specific studies in the following chapter.

II

TESTS OF BASIC SKILLS: AN ASSUMPTION AND A STRATEGY

The Basic Skills Test (BST) Package described below was designed as a research vehicle which would facilitate exploratory research. The test package has potential as a diagnostic tool, and could be expanded into an instructional program. At present, it serves as a matrix by which miniature experiments are conducted in areas of relevance to beginning reading instruction. We have thought it premature to concern ourselves with formal aspects of test construction and evaluation until more has been learned about the significance of specific component skills and the adequacy of alternative methods of assessment.

The principal assumption guiding this research is the existence of separable and independent performance skills which are prerequisite to the acquisition of literacy. The assumption of specific and independent skills is implicit in much that has been written about reading, although the implications of the assumption have rarely been realized in practice. Stern's (1968) report on the UCLA Preschool Language Project is a recent example of this approach, or one can go back four decades to Monroe (1932) who conducted similar research most competently.

An opposing point of view might propose requisite abilities of a more global or maturational nature, as implied by concepts such as "reading readiness" or "critical age." Under the assumption of specific skills, one is led to try to identify specific deficiencies which are then the focus of remedial procedures, while the assumption of a global factor would suggest the adoption of testing and training procedures of an amorphous type.[1] This contrast can be seen in the comparison between the highly structured preschool program of Bereiter and Engelmann (1966) and enrichment or language-experience programs such as those of Alpern (1966) or Brottman (1968).

Starting from the assumption of separable component skills, we have adopted the following strategy in identification and investigation of specific skills. To start, the skill area is delimited as precisely as possible. Then subtests are designed which "surround" the skill, the aim being to select convergent subtests which sample the skill in a variety of different contexts (Haber, 1969). For example, in tests of visual matching ability described below, several variations in task format and materials were employed. Single-letter matching tests were administered in simultaneous and successive modes (i.e., the letter to be matched was either present or absent when the test letters were shown), and stimulus materials were varied over a wide range, from real toys which the child could actually manipulate, to groups of letters printed in upper or lower case.

An attempt is then made to identify the underlying psychological processes involved in the ability. It is at this point that the "how" of testing becomes a major concern. The clarity of the instructions, the quality of the stimulus materials, the skill of the tester—these and similar general factors always contribute to a greater or lesser extent to a child's performance. Ideally, the contribution of these general factors should be minimal.

Our approach to this problem has been to establish a basal level of performance by starting with materials of minimum difficulty for the skill being tested. If the error rate is too high with these materials, we conclude that the testing procedure is at fault and must be improved. If the basal performance is satisfactory, the test is systematically complicated by variation in materials or task procedure until the children begin to make errors. Then

[1] A more or less equivalent form of the global hypothesis states that reading performance depends on a single primary factor such as visual perception, auditory-visual coordination, "language development," etc.

we look for "miniature" training procedures which can be introduced into the testing sequence at the appropriate point to help a child over the hurdles.

It should not be thought that the approach is to "try everything and see what works." The original decision to concentrate on the decoding process has meant that priority be given to investigation of visual and auditory-phonetic processes, and of learning spoken responses to abstract symbols. Some areas (e.g., vocabulary) have been selected for study on the basis of our judgment and evaluation of the work of other investigators.

In summary, the BST package was designed as an in-depth test of a limited set of cognitive skills. Although constructed in the form of a diagnostic test, the package may also be viewed as a matrix within which miniature experiments were conducted. Many of the experiments were duds, in the sense that the basal performance was so poor as to be of no use in isolating sources of difficulty. From our point of view, such results mean that work remains to be done to simplify the testing procedure. These tests aim primarily at uncovering psychological processes, and only secondarily at prediction of later achievement. Intercorrelations between subtests will occasionally be reported, mainly to demonstrate the extent of independence of the measures.

III

BASIC SKILLS TESTS: GENERAL METHOD, SUBJECT SAMPLES

The BST package covers five areas of cognitive functioning: matching of visual forms, auditory-phonetic identification, letter-sound association, vocabulary knowledge, and general achievement. First the general procedures used in administering the test package will be described; then the findings within each of these areas will be presented and discussed. The entire series of subtests is listed in Table 1. The series was divided into three sets, each of 30 to 50 minutes' duration, and given at 1-month intervals between November 1968 and January 1969 to kindergartners in Madison, Wisconsin. This group of children will be referred to as Msn. In March 1969, 10 of the subtests in Table 1 were administered in a single session to kindergartners from a school in Beloit, Wisconsin. This group will be referred to as Bel. The Bel group was tested to replicate some of the more prominent findings from the earlier testing.

The Msn sample consisted of 21 subjects, 11 boys and 10 girls, from an afternoon kindergarten class in a predominantly lower-middle class area of the city. The mean age of the 21 Ss in November was 64 months, ranging from 59 to 70 months. The Metropolitan Readiness Test was administered in March 1969. The Bel sample consisted of 22 kindergartners from a predominantly lower-class area. Eleven kindergartners were drawn at random from the morning session and 11 from the afternoon session. The sample included 10 boys and 12 girls, with a mean age in March of 69 months, ranging from 63 to 75 months.[2] The Metropolitan Readiness Test was administered to the class in May, 1969.

Test variety and total test time were considered in arranging tests for the three different sessions administered to the Msn group and in the selection and arrangement of tests for the single Bel session. Within each session, easy and difficult tasks were alternated insofar as possible to sustain attention. The experimenters were male and female graduate and senior undergraduates at the University of Wisconsin, Madison. Detailed scripts were prepared for each of the subtests, and experimenters rehearsed the scripts before giving any of the tests to insure the equivalence of procedures and instructions. Each experimenter tested at least two children for practice. The data from these practice subjects were not included in the results reported below.[3]

Two forms of each subtest were constructed. In some instances, the alternate forms actually comprised different experimental treatments. Children were randomly assigned to form A or B independently for each session with the constraint that approximately equal numbers of subjects were tested on each form. Subjects were tested individually in small rooms at each school. Responses were recorded as unobtrusively as possible. Tests involving spoken responses were recorded with a Uher 5000 tape recorder and a Shure 545L lavalier microphone.[4]

---

[2] Five of the 22 children in the sample were Negro, reflecting the proportion of Negro children in the school. The data from these children have not been analyzed separately.

[3] A few remarks might be made about testing young children over an extended period of time. First, instructions are always crucial, but especially when success in a test depends upon comprehension of key words or expressions such as "rhyme" or "sound the same." The tester must then rely on indirect tests, examples, gestures, or pre-training procedures for

[4] Phonemes will be indicated in slashes / / by the International Phonetic Alphabet.

Table 1

Basic Skills Test Package, Order of Tests, and Time per Test

Test Name	Session/Session Order		Approximate Time (Min.)	Notes
	Msn	Bel		
Visual Matching:				
Matching	I-1	1	5	
Matching-Retest	I-9	8	4-6	One group tested after VE training, one group after Repr. Training
Oddity-Selection	I-5		3	
Memory-Matching	III-2		10	
Alphabet Knowledge:				
Production	I-2	2	1	
Recognition	I-3	3	2	
Acoustic-Phonetic:				
Rhyming	I-6		6	Same-different
Rhyming-Retest	II-6		4	Same-different
Initial-Sounds	I-10		6	Same-different
Rhyme-Production	II-5	5	3	
Segmentation	III-3	9	10	
Letter-Sound Association:				
Context-Learning	I-11		7	
Letter-Word	II-4		10	
Alphabet-Learning	III-1		15	
Vocabulary Knowledge:				
Picture-Naming	I-4	4	8	
Line-Drawing-Naming	II-1	7	6	
Sorting	II-2		10	
General Abilities:				
Word-Memory Span	I-7	6	2	
Simple Directions	I-8 II-3		3	"Simon-Says" game

Note—Msn sample was tested over three sessions, I, II, and III in November, December, and January, respectively; Bel sample was tested in a single session in March.

communication. Second, for a test session of half an hour or more, keeping the child's attention is a serious problem. In our experience, attention is best maintained if verbal instructions are kept to a minimum. Constant changes in the format of the tests helps, as does alternation of easy and difficult problems.

IV
VISUAL TASKS

To read English a child must learn to isolate, differentiate, and identify the letters of the alphabet. He must learn both to identify these symbols in spite of variations in type styles and size, and to disregard structural variations such as upper- and lower-case forms. He must learn to work with ordered sequences of these symbols, noting the elements in such sequences, and treating as equivalent only sequences in which both elements and order are the same.

METHOD

The visual tasks in the BST package were designed to evaluate relatively low-level skills in this area, using multiple-choice or recognition responses; identification, labeling, and reproduction performance were not measured. Task, instruction, training, and test materials were varied over the several subtests.

On Matching and Matching-Retest, the child was shown a standard stimulus together with a test set of four or five alternatives, and was asked to pick the single alternative that matched the standard. In Memory-Matching, the standard was presented alone, removed, and then the test set was presented 2 seconds later. The child thus had to rely on his memory of the standard in making the match. Finally, in Oddity-Selection the child had to point out the odd member in a set of four alternatives.

Materials were varied in the following ways. To determine the difficulty of the matching task itself, small, three-dimensional toy objects were used as the first few items of Matching. The remaining items were printed in capital and lower-case letters, first single letters and then letter groups—pairs, triples, and quadruples. The items were printed on strips of tagboard in a horizontal row of letters .5 in. high, with the standard on the left. The test sets always included one or more alternatives designed to be highly distracting. Thus single-letter and letter-pair standards were matched with visually confusing alternatives (e.g., b and d). For letter-pair standards, the test always contained an order reversal (e.g., pq for qp). For groups of three or four letters, the test set consisted solely of permutations of the letters in the standard (e.g., rmn, mrn, nmr). The odd member in Oddity-Selection was either visually similar (single letters), or a reversal (letter pairs), or was formed by a permutation of the letter sequence or replacement of a single-letter in the sequence by a visually similar letter (three or four letters). An example would be VWV, VVW, VWV, VWV.

Preliminary analyses of matching data from pilot studies had revealed that most errors involved a failure to preserve order (e.g., if SZ was the standard, ZS was as likely to be selected as SZ). Accordingly, prior to the Matching-Retest, children were given one of two types of supplementary instructions or pretraining. Either the importance of order was verbally elaborated upon using a single example (V. E.), or the child received a series of seven items in which he arranged individual letters to match a standard, with feedback on the correctness of the reproduction (Repr.).

RESULTS

Single-Letter Items

Performance on the various subtests is summarized in Table 2.

The effects of variation in material are quite apparent. Only the Msn sample was tested on the toy or "object" materials. On the first two items, the set of alternatives consisted of four very different toys. There were 2 errors out of 42 possible, an error rate of less than 5%. On the third "object" item, the standard was a safety pin and the alternatives were three safety

Table 2

Percentage Correct Responses on Visual Matching, Related Subtests of BST, and Metropolitan Readiness Matching Subtest

Test/Materials	Msn		Bel	
	Mean	S.D.	Mean	S.D.
Matching:				
Objects (3)[a]	89	19	--	--
Single Letters (10)	83	12	87	11
Double Letters (5)	36	36	48	28
3-4 Letters (7)	30	16	35	54
Matching-Retest:				
Double Letters (3)	67	26	56	39
3-4 Letters (6)	40	28	33	71
Memory-Matching:				
Single Letters (4)	93	24	--	--
Double Letters (5)	42	20	--	--
3-4 Letters (7)	42	25	--	--
Oddity-Selection:				
Single Letters (2)	86	23	--	--
Double Letters (4)	29	25	--	--
3-4 Letters (6)	29	23	--	--
Alphabet-Production:	27	27	33	31
Alphabet-Recognition:	23	77	18	82
Metropolitan Readiness:				
Matching (14)	69	19	49	25

[a]Number in parentheses is number of different items included in each category. Guessing rate is 25% except for Alphabet Tests (4%), Metropolitan (33%), and 8 of the 10 single-letter matching items (20%).

pins of varying size and color, and a red whistle. Increasing the similarity of the test set increased the error rate to 25%, indicating the importance of the make-up of the test set.

On the other hand, analysis of the single-letter items indicates that for our samples of children, the letters in the English alphabet were not perceived as confusing, though they had been selected as such. Except for the first two items, the single-letter series were constructed to be quite difficult, since pilot testing during the summer of 1968 had produced extremely low error rates on single letters. Typical materials were G - C Q G D O, b - h d b f k, and r - n m w r u. "Difficulty" was based in part on feature analysis of the upper- and lower-case letters (Gibson, 1965), and in part on findings that right-left and up-down reversals often produce confusions.

The right-left mirror image transformation was the major source of errors. Three single-letter test sets included a right-left reversal of the standard (e.g., b d). The error rates on these items were 46% (Msn) and 41% (Bel), with selection of the reversal alternative accounting for 85% of these errors. For the remaining seven items in the single-letter series there were two errors in both the Msn and Bel samples out of 147 and 154 opportunities respectively. Dunn-Rankin (1968) asked second and third graders to judge the similarity of lower-case letters; for the two sample sets above (b and r standards), the perceived similarity of the test sets is about equal from Dunn-Rankin's ratings, but the error rate for the b item was 40%, whereas the r item was correctly matched 100% of the time.

Letter Groups

The error rate increased sharply when letter groups were tested, particularly triples and quadruples where performance was nearly at chance performance level of 75%. For bigrams, one of the alternatives was an order reversal of the standard pair and two were visually confusable with the standard. Thus, a typical item might be CQ—OQ QC CQ CO. Of the bigram errors on Matching, 70% (Msn) and 65% (Bel) were choices of the reverse order alternative; if the errors had been the result of random selections among the alternatives, only 33% of the errors should have been reversal errors.

In the Matching-Retest, 52% (Msn) and 43% (Bel) of the bigram errors were reversal errors. Unfortunately, the placement of reversal alternatives was not balanced for serial position between the A and B groups, and so a breakdown of reversal errors by treatment groups is meaningless. On Memory-Matching, 54% (Msn) of the bigram errors were reversal errors.

Serial Positions

The items were displayed in a horizontal array with the standard to the left. There were enough errors in the letter-group series for analysis of serial position effects, which showed (Table 3) that in the Matching and Matching-Retest tasks, the children were scanning from left to right. In the Memory-Matching task, no pattern to the errors was evident. In the letter-group items, all of the alternatives were highly similar, in the sense that they were constructed of the same letters as the standard. If letter-order information were not preserved by the child, and if the child scanned the alternatives from left to right, one would predict the distribution of errors would resemble the data in Table 3. The presence of the standard to the left of the alternatives was important in directing the scanning process, since errors were randomly distributed when the standard was not present.

The data on the right of Table 3 provide further information on the scanning process. The percentage of reversal errors in bigram matching are shown as a function of whether the correct alternative was to the left of the reversal alternative (C-R), or to the right (R-C). These data are shown for each of the matching tasks, separately for the V. E. (verbal elaboration) and Repr. (reproduction) groups. In Matching, if the reversal alternative were to the left of the standard in the test set, it was quite likely to be chosen. This result also held in Matching-Retest following V. E. but much less so following Repr. Thus, the training procedure affected the type of errors made by the children, though not the rate of errors. In Memory-Matching, the reversal alternative was equally likely to be selected whether to the right or left of the correct alternative, supporting the conclusion that scanning was unsystematic in this task.

Oddity-Selection

Overall performance level on the Oddity-Selection task, in which the child was asked to point out "which one of these is not like the others" was comparable to performance of the other matching tasks, which would suggest that these tasks all tapped similar component skills. On the other hand, one child insisted that for eight of the twelve items, all four alternatives were identical; in seven other instances a child refused to point out any of the alternatives as different from the others. In the other matching tasks, there were only two cases of this sort. There was no evidence of serial position effects in this task.

Metropolitan Readiness Match

Performance on the Matching subtest of the Metropolitan Readiness Tests (Hildreth, Griffiths, & McGauvran, 1964) was somewhat higher than on the BST matching tests. This subtest consists of fourteen items, with the same format as Matching in BST (horizontal array, standard to the left). A variety of

Table 3

Breakdown of Total Errors, Showing
Position Effects in Visual Matching Tests

Test/School	Per Cent of Total Errors at Each Serial Position				Per Cent Bigram Reversals[a]		
	1	2	3	4	Training Group	C-R	R-C
Matching:							
Msn	63	21	8	8	V.E.	15	67
					Repr.	14	76
Bel	52	23	12	13	V.E.	14	52
					Repr.	14	40
Matching-Retest:							
Msn	78	10	6	5	V.E.	15	70
					Repr.	14	27
Bel	50	27	7	16	V.E.	14	55
					Repr.	5	18
Memory-Matching:							
Msn	21	27	23	29	---	29	33

[a] Per cent of reversal errors on two-letter stimuli when correct choice was to left (C-R) or right (R-C) of reversal choice, by V.E.-Repro. training groups; training given just prior to Matching-Retest.

standard items is used in this test—words from four to nine letters long in upper- and lower-case, as well as various kinds of abstract figures. Each test set contains three alternatives which would produce a higher guessing rate than for the BST matching tests. The alternatives are similar to the standard, consisting of letter permutations for the words and minor figural variations for the forms. Also in BST Matching, the leftmost alternative in the test set was correct for only two of the twelve letter groups, whereas in the Metropolitan Matching, the leftmost alternative is correct for five of fourteen items. Given the previously noted bias for the leftmost alternative, these arrangements would also produce fewer errors on the Metropolitan compared to the BST. Finally, the BST was administered in the fall, the Metropolitan the following spring. Hence, no significance is attached to the higher performance on the Metropolitan Matching test. Comparisons between performance on these various tests will be discussed in a later section on intertest correlations.

DISCUSSION

There are three general considerations in the relation of visual perception to reading. First is the question of isolation, discrimination, and identification of the abstract characters which serve as letters in the English alphabet. Second is confusion of left-right mirror images, which is related to the first question but warrants discussion as a separate issue. Third is the perception, analysis, and storage of written words.

Letter Perception

Most children in this country encounter printed materials at an early age. Few of the kindergartners we tested "knew their ABC's" but most attempted the letter-matching task without hesitation and performed quite well, given the confusability of the English alphabet and the fact that they could not label many letters. Whether matching is conceived as a process involving templates or distinctive-features (Neisser, 1967), the single-letter test sets were highly similar, and are judged as such by second graders (Dunn-Rankin, 1968), and yet error rates were quite low except for right-left reversals.

Gibson, Gibson, Pick, and Osser (1962), in what has become a classic study, investigated the degree to which letter-like standards were confused with various transformations of the standard—perspective, rotation and reversal, line to curve, and break-and-close. The format of their test was generally the same as that used in BST. There were 12 alternatives in each test set, and the children were instructed to select as many alternatives as they wished from the set, since sometimes more than one copy of the standard would be included.

From the ages of 4 to 6, both right-left and up-down reversals produced errors (choice of a transform as equivalent to standard) at a rate of from 50% to 15%. Kindergartners were also tested on transformations of real English letters and here the error rate for right-left reversals (b, d) was more than double that of up-down reversals (b, p)—19% vs. 8.5%.

Any attempt to identify transformational relations among real letters is risky, but as an example, if it is assumed that C and O are related by a break-and-close transformation, or V and U by a line-to-curve, then the error rates in Gibson, et al., would lead one to expect much higher error rates in BST matching tests than were actually found. In both of our samples, the overwhelming preponderance of errors made by the children on single-letter tests were right-left reversals; only 10 out of 70 errors could be classified as up-down reversal, close-break, or line-to-curve.

The substantial similarities among English letters are indicated by the difficulty in devising reliable pattern-recognition devices (Uhr, 1966). The Gibson, et al., investigation suggests that transformational relations may be a useful way of describing perceptual similarity. What remains to be formulated is a model of stimulus structure which can complement feature-transformation analyses. Gibson's (1965) approach took the form of a "feature listing," in which, for example, Z was equal to <straight segment, horizontal, oblique/; discontinuity, horizontal>. This listing provides all the elements or features in Z, but lacks a description of how the elements are structurally related. A more complete list would include features and rules of combination. Thus p might be roughly described as <straight segment, vertical, descender; curve, closed>. This list consists of writing instructions which are applied from left to right, with semicolons denoting operational breaks.

The point being made is just this: the perceptual similarity of two letters depends not only on the transformational relations between the objects but also on the structure of the objects. We know that d and b are easily confused by kindergartners, and we have to live with the fact that both are letters in English—but what about S and ƨ, 5 and ꙅ, or N and И? Is it true in general that the right-left transformation is a source of confusion, or does this hold only for certain structural classes of stimuli? An answer to this question might lead to a better understanding of the nature of b-d and p-q errors.

The matching task has much in common with visual search tasks of the sort investigated by Neisser (1967), in which subjects are presented one or more items to remember, and are then asked to look through a list until they find one of the memorized items. Instructions stress speed—the target is to be reported as quickly as possible. Neisser has shown that increasing visual similarity slows down search rates in college students, and Gibson and Yonas (1966) have reported similar results using subjects as young as second graders.

The task of reading itself also involves heavy memory and visual loads; and although speed may not be overtly emphasized, the laggard is certainly penalized. There is little research on the effects of item complexity, memory load, visual load, visual similarity, and stress of speed vs. accuracy on matching performance. These variables denote, in turn: the choice of letters, words, familiar or nonsense forms as item material; the number of alternatives being searched through; the similarity (transformational or featural) of the alternatives in the test set; and the degree to which instructions require the subject to respond as quickly as possible or as accurately as possible. The Gibson and Yonas study suggests that in speeded search, visual similarity may affect performance. With similar materials but minimal memory and visual load, and minimal stress on speed, performance was essentially perfect in the BST matching tests of single letters. The difficulty in the more demanding search task would appear to be cognitive rather than perceptual, except possibly for the problem of right-left reversal confusions, to which we will now turn our attention.

Right-Left Reversals

The fact that children confuse the letter pairs b-d and p-q has been repeatedly documented (see Fellows, 1968, and Benton, 1959, for reviews). The problem is universal, profound, and persistent. Substantially more than half of all kindergarten children confuse these pairs sometimes. In a two-choice test, error rates of 15-25% are common. The difficulty may persist until 9 or 10 years of age in the case of children who are "nonreaders." The problem has been attributed to various kinds of dysfunction—physiological, perceptual, memorial, or cognitive. Various training procedures have been evaluated, with mixed results (Strang, 1967; Harris, 1969; Jeffrey, 1958). There have been no tests of the long-term effects of such training on reading or reading-related tasks. The degree of confusion appears to depend on the physical arrangement of the stimulus objects. For example, Huttenlocher (1967a, 1967b) presents evidence that right-left reversals occur in a horizontally arranged display, but when a vertical array is used, up-down reversals are more frequent than right-left reversals.

If it were not for the existence in the English alphabet of the pairs b-d and p-q, this entire discussion would be of interest only to investigators of spatial orientation (Howard and Templeton, 1966). Various efforts have been made to determine the number of reading errors due to right-left reversals; this is probably a hopeless task. It is clear that children do make such errors, even under optimal testing conditions, and that we do not understand the sources of such errors or how to remedy the problem. One simple solution might be to use only upper-case letters in initial reading instruction, or to adopt some sort of stylistic variation in the typography used in initial reading, (i.e., replace d with ∂ and p with ρ).

Word Matching

Next we will consider matching of letter groups. Although there is considerable evidence suggesting that experienced readers perceive words as units rather than as letter strings, the manner in which beginning readers or prereaders process words is less certain. Surprisingly, there does not appear to have been a test of the simplest hypothesis, viz., that the probability of error in matching a standard word with a particular test alternative is the product of the error probabilities for the constituent letter pairs—assuming independence of the letter comparison processes. The problem is an interesting one, because there are good arguments to suggest that error rates in word matching should be greater or less than predicted by independent scanning: greater because of the increased information-processing load; less, because word configuration cues are another source of information. Although there is anecdotal evidence to the effect that word configuration may be important in scanning by experienced readers, there appears to be little research on the effects of configuration on matching by beginners. (Our data do not provide a sufficient range of tests of single or multiple letter groups for such a test.)

An important source of word matching errors in the BST data and other studies involves confusions between a standard and an alternative consisting of an order permutation of the letters in the standard. A special case of this situation is the right-left reversal of a word—was for saw. Two of the better studies on word perception in prereaders come from the older literature. Hill (1936) investigated word-matching ability in kindergartners and first graders with a well-constructed test designed to evaluate the effectiveness of cues at different locations within a word. She found that the first and last letters were more salient than those in the middle; that the error rate was highest when an alternative preserved the configuration of the standard; and that the highest error rate was observed when two words differed by a single mirror-image letter (e.g., rimd vs. rimb). Hill did not test mirror-image word alternatives, but from her data it is clear that other sources of confusion contribute substantially to word matching errors. First graders made fewer errors than kindergartners, but the pattern of errors was the same in both groups.

Davidson (1934) compared form and word reversal confusions in kindergartners and first graders. A five-alternative multiple-choice test was employed, the child being instructed to select one or more alternatives that were the same as the standard. For the "nonsense" geometric forms, one of the alternatives was a right-left reversal; for the words, one of the alternatives contained the letters of the standard in reverse order. Confusions were common with both types of materials; 94% of the kindergartners and 62% of the first graders made at least one form reversal, and the corresponding figures for words were 83% and 33%. The percentage of errors, reversal and others, for forms and words, is rather interesting (Table 4). Reversals account for more than two-thirds of all form errors in both kindergarten and first grade children, whereas "other" errors are in the majority in the word tests. [Davidson describes the other alternatives simply as "words known to be easily confused with the key word."] Thus, Davidson's results suggest that order reversal errors are not necessarily the singular source of confusion in word perception that mirror-image reversals are in single-letter matching.

Table 4

Error Distributions in Visual Matching of Forms and Words by Kindergarteners and First Graders (Davidson, 1936)

Grade	Forms (10)[a]			Words (15)[a]		
	Reversals	Other Errors	Total	Reversals	Other Errors	Total
Kindergarten (n = 50)	23	11	34	22	26	48
First Grade (n = 120)	13	5	18	5	8	13

[a]Number of items in test.

Another relevant study is that of Pufall and Furth (1966). In one portion of that study, children from 4 to 6 years old classified pairs of items as same or different. The items consisted of triads of two objects of one kind (A) and a third object of a different (B) (e.g., AAB or ABA). On each trial, the child was presented with a pair of triads which had the same or different sequences (e.g., AAB - AAB or AAB - BAA). Various types of materials were used—colored marbles, colored cards, familiar geometric forms, line drawings of familiar objects, and nonsense forms. Given the high intralist similarity, error rates were relatively low, ranging from 29% at age 4 to 0% at age 6. The result most pertinent to order errors was the finding that 55% of all errors were mirror-image reversals; i.e., AAB and BAA were identified as "same" more than half the time. When the memory component of the task was increased by presenting the two triads successively rather than simultaneously, mirror-image reversal errors occurred at about the same rate, but the rate of other types of errors (e.g., AAB - ABA, "same") doubled.

In summary, word matching was a relatively difficult task for kindergartners, even though they were able to match the component letters. Words were handled as strings of more or less independent letters, although confusions due to configurational similarity also existed. Children did not appear to match solely on the basis of single-letter identity, although initial and final letters were more likely to be correctly matched than middle letters. Permutation confusions are common and although there is little systematic research on the problem, right-left order reversal confusions appear to occur frequently in word matching but not so exclusively as right-left mirror-image confusions in single-letter matching.

V
ACOUSTIC-PHONETIC RECOGNITION

We have argued elsewhere that if a child is to learn to read English, almost certainly he must learn to make use of letter-sound correspondences (Venezky, Calfee, & Chapman, 1969). Learning these correspondences involves analysis and identification of strings of letters and of the phonetic units which correspond to letters. Most literate adults unquestioningly accept a "discrete-units" theory of phonology, probably because they confuse letters and sounds. Yet research on this problem to be discussed below suggests that prereaders have considerable difficulty performing many tasks which require analysis of a spoken word into phonetic components.

The acoustic-phonetic subtests in BST employed a variety of procedures to evaluate children's competence in this area—recognition of word pairs having the same or different initial or final phonological segments, production of rhymes, and acquisition and transfer tests on a paired-associate problem which could be "solved" by noticing that the responses were phonological segments of the stimuli (e.g., FEEL-EEL).

METHOD

Three of the tests were same-different tasks —Rhyming, Rhyming-Retest, and Initial-Sounds. In each, the child was informed by instructions or pretraining that certain features of word pairs would be critical, the final -VC segment in the rhyming tests and the initial C- segment in Initial-Sounds. The child then was asked to respond "Yes" or "No" to a series of word pairs, depending on whether the critical feature was the same or different in the two words.

In Rhyming, the child was asked whether or not two words "sounded the same at the end"; a pretraining series with corrective feedback was used to exemplify this concept. The Rhyming-Retest was the same as Rhyming except for materials, but was preceded by the Rhyme-Production task to be described shortly. In Initial-Sounds the child was asked whether or not "two words start with the same sound." Three continuants were tested, /s/, /m/, and /l/. To simplify the task for the child, all tests of a given phoneme were administered in a block. Two different testing procedures were used. Children in Group A received two exemplars of the critical phoneme at the beginning of each series. Then a series of test items was presented and the child was asked whether or not each one began with the critical sound. Children in Group B were also required to repeat one of the exemplars together with the target word before the test question was asked.

The Rhyme-Production test, designed as another means of testing rhyming ability, consisted of two parts. First, the experimenter gave two rhymes for each of eight words spoken by the child in a picture-naming task. Thus, if the child said "girl" when shown a picture, the experimenter might say "pearl, curl." Then the child was asked to switch roles, to give a rhyme for each word pronounced by the experimenter as the name of a picture. The experimenter told the child he was right if the child produced a rhyme, and gave two rhymes if the child was wrong or gave no response.

The Segmentation test consisted of two paired-associate tasks, each task consisting of five study-test trials on a list of three pairs followed by a transfer test on six different but related stimuli. The stimuli were all familiar CVC words, and the responses were the final -VC segment of each word. For one list (RW), the responses both in training and transfer were real words: [F(EEL), SH(OUT), P(ILE)], and for the other list (NS) the responses were nonsense: [S(OAP), R(IDE), CH(IEF)].[5] Children in Group A received the NS list first followed by the RW list. Group B received the lists in the reverse order. The children were urged to give a response on every test, even if it was a guess.

RESULTS

The Rhyming, Rhyming-Retest and Initial-Sounds test results are quickly described. In each test, the percentage of correct responses was 49%, compared with a guessing rate of 50%. In short, none of the children did better than chance on any of the tests. Some variation in performance was apparent in response biases but it is clear that the children tested could not perform any of these tasks—because of an inability to grasp the concepts of phonetic matching, because the instructions or pretraining were inadequate or other short-comings of the "same-different" procedure, or any combination of these reasons.

The Rhyming-Production test produced the most interesting results. The percentage of rhymes produced were 39% (Msn) and 37% (Bel). Some children in both groups refused to attempt the task at all (5/21 (Msn) and 4/22 (Bel)); for those children who made an attempt, failures to respond were 9% (Msn) and 26% (Bel). Some children were quite good at the task; 10/21 (Msn) and 10/22 (Bel) produced four or more appropriate rhymes in eight attempts. Except for three children who produced a single correct rhyme, the remaining children failed on all eight items, producing noticeably bimodal distributions in both samples. A substantial percentage of the rhymes were nonsensical, 51% (Msn) and 44% (Bel), indicating a grasp of the phonological nature of the task. Only a very few of the responses were semantic associates (e.g., pot - coffee cup).

In Table 5 are the results of analyses of the Segmentation test in which children learned two paired-associate lists based on phonological relations between the stimulus and response members and were tested for transfer of this learning. Analysis of variance was carried out on several dependent variables: (a) number of correct responses, (b) phonological errors, (c) stimulus repetitions, and (d) other errors. The independent variables were school, group, RW vs. NS, and training vs. transfer. Aside from confirming that performance of one of the Bel groups was significantly poorer than the other three groups, the most interesting result was a significant ($p < .05$) interaction between RW and NS materials during training and transfer. The details of this interaction are displayed in Fig. 1. During both training and transfer, about .6 of the responses were either correct or phonological errors. (Phono-

[5] The data do not provide an especially strong test of the role of response familiarity, because the responses in the RW list, (-OUT, -EEL, -ILE or "aisle"), were not much more familiar to the children (except for -OUT) than the NS responses (-OPE, -IEF, -IDE).

Table 5

Percentage of Correct and Phonological Responses on Segmentation Test During Training and Transfer

Trial	Training					Transfer
	1	2	3	4	5	
Real Word:						
Correct	35	44	50	51	60	24
Phonological Error	21	15	10	9	9	28
Nonsense:						
Correct	27	34	44	44	48	24
Phonological Error	25	23	19	19	13	27

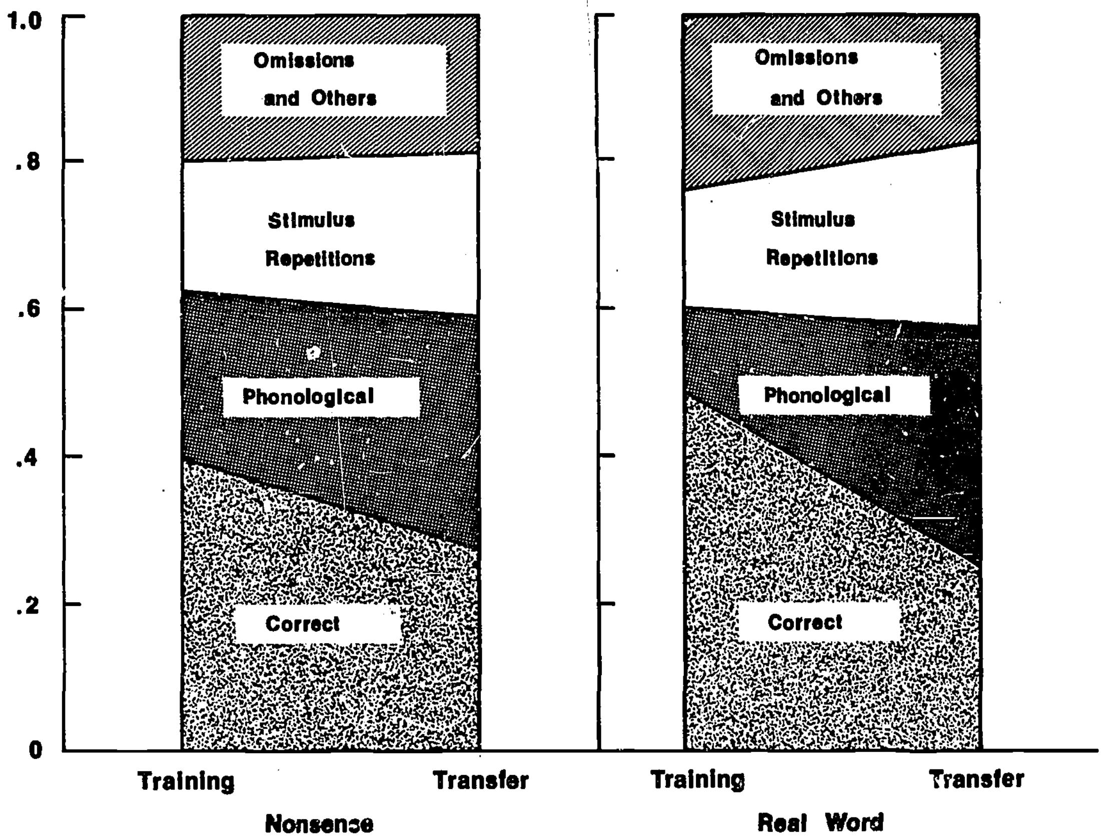

Fig. 1 Segmentation Responses to Real Word and Nonsense Lists on Training and Transfer

logical errors were defined as monosyllabic responses in which the vowel matched that of the stimulus word, excluding repetitions of the stimulus.) In transfer these two classes of responses occurred about equally often. In training on RW lists, .8 of these two responses were correct but only .65 of these two classes of responses were correct in NS lists. In other words, RW and NS lists produced equal numbers of phonologically related responses in both training and test; RW lists produced a larger proportion of exactly "correct" responses but this was not significantly related to transfer performance.

There were few misplaced responses (e.g., SHOUT - EEL) in either the RW or NS list, except for the response OUT which was given incorrectly in the RW list a total of 34 times (out of a possible 430 opportunities). There were 9 other misplaced responses in RW lists and a total of 8 misplaced responses in NS lists.

To the extent that the children learn a concept based on the phonological relations between the pairs, one might expect to observe the following: (a) transfer to a new list of stimuli for children who reached criterion on the training list, (b) more rapid learning of the second list compared to the first, (c) errors should be phonologically related to the stimulus word, and (d) the distribution of correct responses should be somewhat bimodal in both training and transfer.

Except for (b) which might be predicted on other grounds as well, each of these hypotheses found support in the data, indicating that many of the children were sensitive to the phonological character of the pairs. Several children responded with an initial consonant substitution rather than a deletion as required by the task. For example, one child produced WINE, WEAM, WOSS, and WAFE for the transfer items PINE, TEAM, MOSS, and SAFE, respectively.

The bimodal character of the distributions of correct responses in training and transfer is noteworthy. If a child had grasped the phonological nature of the task by the end of training, most of the responses on the last training trial and on the transfer trial should be correct responses or phonological errors. A strictly all-or-none concept model would postulate that if a subject had learned the phonological concept, only phonological responses would be produced at the end of the training and during transfer; otherwise such responses would occur at a chance rate. This model is certainly too strict; there would be some correct responses due solely to rote learning and some errors due to attentional lapses even though a child had grasped the concept. The distributions in Table 6, showing the number of children making various numbers of correct or phonologically related responses on the last trial and during transfer, are significantly bimodal when compared, for example, with a binomial distribution. Some children achieved perfect or near perfect scores, and most others made no correct responses, or at most one. Few subjects performed at an intermediate level.

Some children (3 in Msn, 4 in Bel) persisted in simply repeating the stimulus word. The study-test procedure which has proven an efficient means of teaching paired-associate lists to college students (Battig, 1965) was inadequate in interrupting stimulus repetition errors, perhaps because it did not provide immediate negative feedback to the children. One child (in the Bel sample) responded to each stimulus word with another word (e.g., BOOK, HORSE, CAT) as though performing a word-association task. Except for the eight subjects just described, there were very few "real word" responses which were not phonologically related to the stimulus word.

In summary, although on the Rhyming, Rhyming-Retest, and Initial-Sounds children were unable to deal with phonetic segments by identifying words as same or different on the basis of such segments, on Rhyme-Production and Segmentation children produced words or word-like units based on phonological segmentation of stimuli. Moreover, there was a definite relation between a child's performance on the Segmentation and Rhyme-Production tests, as shown in the contingency table in Table 7. Also shown in Table 7 are the contingency tables for Segmentation and Alphabet-Production, and Segmentation and Alphabet-Recognition. This result suggests that these two tests are convergent: that both are tapping the same set of acoustic-phonetic skills. The relation between segmentation and the alphabet knowledge tests, weaker than the previous relation but statistically significant, could be taken as evidence that acoustic-phonetic skills are important in later reading achievement, or as a measure of a general performance factor in all the tests, or both.

The possibility that we may have begun to isolate acoustic-phonetic skills in prereaders is encouraging. There remains the problem of relating this skill to other tasks requiring analysis of word sounds, particularly the isolation and identification of phonemes, and of

Table 6

Distribution of Subjects on Segmentation Test; Number of Subjects Making 0, 1, 2, or 3 Correct or Phonological Responses on Final Training Trial, and 0, 1, . . . 6 Such Responses on Transfer Trial

	Number Correct or Phonological Responses						
	0	1	2	3	4	5	6
Final Training Trial:							
Real Word	6	10	3	24	--	--	--
Nonsense	8	6	5	24	--	--	--
Transfer Trial:							
Real Word	6	9	2	1	3	5	17
Nonsense	13	2	0	3	3	5	17

Table 7

Contingency Tables Showing Relation of Performance on Segmentation Test with Rhyme-Production, Alphabet-Production, and Alphabet-Recognition Tests
Children in each Sample Partitioned into Those Above and Below Median

Segmentation	Rhyme-Production		Alphabet-Production		Alphabet-Recognition	
	Above	Below	Above	Below	Above	Below
Above	17	5	15	6	14	7
Below	3	18	7	15	7	15
	$X^2 = 17.1$ $p < .001$		$X^2 = 6.7$ $p < .01$		$X^2 = 5.9$ $p < .05$	

devising and evaluating various training procedures which could be used with children who are deficient in this area.

DISCUSSION

There has been relatively little research on acoustic-phonetic skills in children. Even the simplest problems remain to be fully investigated, such as the types of confusions that occur in phoneme perception. Some work has been done on children's discrimination of minimal pairs using a same-different procedure (e.g., "does /ba/ sound the same as /pa/?"; Deutsch, 1967; Templin, 1957). Unfortunately this procedure is difficult for young children, particularly those who are most lacking in general language ability. Studies recently completed in our laboratory (Kamil & Rudegeair, 1969; Skeel, Calfee, & Venezky, 1969) show that under optimal conditions, kindergartners and even younger children could discriminate between minimal phonetic contrasts, even when meaningless CVC's are used as stimuli. One factor in obtaining satisfactory performance levels by pre-readers on this task was avoidance of the same-different test procedure. Another factor was the use of repeated test sessions; in the study by Kamil and Rudegeair, the error rate on the second day of testing was about half that on the first day. Half of the errors which did occur involved phonetic contrasts known to be difficult for adults—/f - θ/ and /v - ð/.

An early study by Bond (1935) is perhaps the most complete and informative study of auditory factors related to reading available. His subjects were second and third graders who were reading at grade level (Control) or were below grade level 6 months or more (Experimental). Half the children had been taught reading by a phonetic method; half, by a look-say method. Six different auditory tests were administered, as well as standardized reading achievement tests. The auditory tests were AA (auditory acuity); SB (sound blending; e.g., tester says /bə - ai/, student says /bai/); AP (auditory perception, a potpourri: (a) giving the letter name for different sounds as in /bi/ for /bə/, (b) giving words beginning with a sound such as /kait/ for /kə/, and (c) giving words ending with a "sound" such as /kip/ for /-ip/); AM (auditory memory measured by digit span and nonsense-syllable span); AR (auditory rhyme, reproduction of various tapping patterns), and AD (auditory discrimination, a same-different test of minimal pairs such as /res - rez/). Finally, an articulation test was given in which responses were elicited by line drawings of familiar objects.

Bond's analyses showed that on SB and AP the Experimental Group was significantly poorer than the Control Group in both the Phonetic and Look-Say conditions. On Test AA the Experimental Group was poorer than the Control in the

Phonetic condition only, while on AM the Experimental Group was poorer than Control in the Look-Say condition only. It is tempting to conclude that to benefit from phonics reading instruction it is important to have good hearing; whereas in whole-word instruction it helps to have a good memory. AD, AR, and articulation ability were not significantly related to reading performance. AR is akin to what has been investigated as "auditory-visual integration," the ability to match auditory-temporal patterns (e.g., a series of dots and dashes as in Morse Code) with a corresponding visual-spatial pattern. Performance on this kind of test has been shown by several investigators to be correlated with reading achievement. The most complete study of this problem (Muehl & Kremenak, 1966) showed that cross-modal matching (auditory-to-visual and visual-to-auditory) was correlated with first-grade reading achievement (Pearson r's of .52 and .39) but intramodal matching (auditory-to-auditory and visual-to-visual) was not. Bond's AR test was essentially auditory-to-auditory and so both sets of data are consistent in saying that intramodal pattern matching is unrelated to reading achievement.

Muehl and Kremenak also combined the four auditory-visual-integration scores with various reading readiness scores in a discriminant function analysis and found that the letter-naming readiness score was the only significant predictor of first grade reading achievement. The predictive value of the child's knowledge of the alphabet in kindergarten is well known. Equally apparent is the diagnostic uselessness of this information. Cross-modal pattern matching ability may be a poorer predictor, but would seem to have greater potential for diagnosis of a specific deficiency.

Two other recent studies have examined acoustic-phonetic segmentation skills in young children. Bruce (1964) had children between the mental ages of 5 and 9 say what would be left if a particular phoneme was deleted. The phoneme was either initial, medial, or final (e.g., BRING, WENT, OR EVERY, the underlined letter corresponding to the phoneme to be deleted). Relatively few children were tested at the younger ages and the stimulus words were quite varied. His pretraining procedures were well designed to guarantee high levels of performance. Nonetheless, children with a mental age of 7 or less were completely unable to perform the task. By age 8, more than half of the responses were correct. Bruce's analysis of errors provides an interesting picture of the development of segmentation ability during the early school years. The details are beyond the scope of this paper but, briefly, from age 6 on, disregarding omissions and stimulus repetitions, more than 90% of the errors are phonologically related to the stimulus word.

McNeil and Stone (1965) first trained kindergartners to hear /s/ and /m/; one group was trained on 24 real words, a second on 24 nonsense words. The criterion test consisted of four real and four nonsense words, for each of which a child was asked, "Do you hear /s/ or /m/ in ________?" Following training on the real words, 58% of the criterion responses were correct compared with a guessing rate of 50%; following nonsense-word training, 78% of the criterion responses were correct. Real words were used exclusively in the same-different tests in the BST package; McNeil and Stone's results suggest that by using semantically anomalous stimuli, children might more easily focus on the phonological properties of the words.

Identification of significant acoustic-phonetic skills is a significant first step, but this must be followed by the development and evaluation of remedial training programs to realize the full potential of this approach. The work of the Russian psychologist Elkonin (1963) is the most interesting effort at training of phonetic or word-analysis skills of which we know. The details of his procedure are a bit obscure but in general this training procedure taught children to match phonemes with blank tokens which they arranged on a formboard with as many spaces as the word had sounds. A picture representing the word was also present. Gradually, specific identification of sounds was introduced and the formboard removed; the child had to select the appropriate number of tokens and, after repeating the word, break it up into individual sounds. In spite of the scant details, Elkonin's technique looks promising and the results are encouraging.

The investigations above suggest that the problems of kindergartners in acoustic-phonetic analysis are not sensori-perceptual —children can discriminate minimal phonetic contrasts about as well as adults and can perform some tasks which require phonetic segmentation. Instead, the problems are cognitive, as Vernon (1960) concluded: ". . . The children (backward readers) were familiar with the general shapes of words and the letters they contained, but their knowledge was quite unsystematic. They knew something about the shapes of letters and letter groups and their associated sounds and that the letters had to be blended together to form the words but they were so confused that they had no certainty as to the correct manner of performing these processes, or of coordinating them together.

"The most common feature of reading disability is the incapacity to perform the cognitive processes of analyzing accurately the visual and auditory structures of words. The backward reader guesses wrong letters, or the right letters in the wrong order . . . .

"The fundamental and basic characteristic of reading disability appears to be cognitive confusion and lack of system. Why even quite an intelligent child should fail to realize that there is a complete and invariable correspondence between printed letter shapes and phonetic units remains a mystery which has not yet been solved. It must be attributed to a failure in analyzing, abstraction, and generalization, but one which, typically, is confined to linguistics . . . (p. 71)."

VI
LETTER-SOUND ASSOCIATION

This section focuses on tasks which required the children to learn to associate sounds and letter-like symbols. These tasks were analogous in many respects to the early stages of reading instruction. The materials were relatively abstract and unfamiliar. Various task formats were investigated; in some tests a spoken response was given to a visual symbol, and in others the experimenter pronounced a stimulus "word," and the child had to pick out the corresponding visual symbol. Most of these tasks were quite difficult and learning was slow. Nonetheless, detailed analysis of the data provided some clues to specific sources of difficulty in these tasks.

METHOD

Alphabet-Learning was analogous to memorizing the names of the letters of the alphabet. Six nonsense forms (after Gibson, et al., 1962) served as stimuli, three paired with a consonant-vowel response (/we/, /ge/, /je/) and three with a vowel-consonant response (/æt/, /æd/, /æb/). These stimuli were chosen to parallel the pattern in the English alphabet represented by such sets as B, D, G and L, N, S. Except for /æb/, all of the responses were real words.

In this and the other learning tasks, study and test trials alternated. In this particular task, on the first half of each trial, three items were studied and then tested, and on the second half of the trial the remaining three items were studied and then tested. The list was always split on each trial to produce a mixed response set; i.e., the /-e/ and /æ-/ subsets were never presented all together. Five trials were administered in this manner.

The Letter-Word task investigated the blending of letter-sound combinations. The Letter list consisted of three unfamiliar letter forms (capital Greek), denoted M, N, and O, paired with the sounds /m/, /n/, and /o/. The Word list was derived from the Letter list by pairing O with M or N to form MO, NO, OM, ON and associated responses /mo/, /no/, /om/, and /on/. Five study-test sequences were given on each list. Half the children learned the Letter list first and then the Word list, the other half were given the lists in the reverse order. In the instructions it was pointed out that the words consisted of two symbols but the individual letter-sound correspondences were not stressed.

The Context-Learning task was directed primarily toward the effect of variability in letter-sound correspondences on learning these correspondences. The test used the capital letters D, C, A, and E as stimuli. First, the child was shown A and E and taught to point to A when the tester said /e/, and to E when /i/ was pronounced. Then the stimulus pairs DA and DE were placed before the child and he was told to point to DA when /de/ was pronounced and to DE when /di/ was pronounced. Then a test was administered; the tester said /di/ and /de/, and the child had to point to either DE or DA. This type of multiple-choice recognition test procedure was used throughout this task. Next, the pairs CA - /ke/ and CE - /si/ were presented for a study-test sequence. [Thus, all of the letter-sound correspondences were invariant except for C which was /k/ before A and /s/ before E.] There followed a series of five study-test trials on the entire set of four pairs, unless a criterion of two perfect test trials in succession was attained. If the child made an incorrect selection, the tester said "no, that makes the sound _____. Let's try again." Finally a transfer test was given; the set DAZ, DEZ, CAZ, and CEZ was presented, and the child was asked to point out in turn "which one says . . ." /dez/, /diz/, /kez/, and /siz/.

Only the Msn sample received these tests.

RESULTS

Alphabet-Learning

Over the five-trial series, the percentage of correct responses to the six-pair list increased from 22% to 33%; the percentage of omissions decreased from 38% to 16%; correct responses as a percentage of responses attempted remained constant at around 38%. In 28% of the attempts, the vowel was appropriate to the stimulus shown. Another 17% of the attempts were intrusion errors from the nonmatching sublist (e.g., /æd/ was given as a response in place of /we/). The remaining errors (17%) could not be classified.

From the figures given above, it might appear that the learning was going on in a reasonable manner. Response omissions were decreasing; more than a third of the attempts were correct; and of those attempts which were errors, almost three-quarters gave evidence of response learning. In other respects the data are discouraging, however. First, all of the improvement took place from the first to the second trial, where omissions dropped from 38% to 19%. From the second through the fifth trials, there was no noteworthy change in any of the measures. Second, not one subject out of twenty-one reached criterion; in fact, only five children made four or more correct responses on the final test of the six-pair list. Third, there was no consistency in the retention of pairs from one trial to the next. For example, of those pairs which were correct on the fourth trial, only 43% were correct on the fifth trial. Comparing this percentage with the unconditional percentage of correct responses on the fifth trial, 33%, it can be seen that whether or not an item was correct on the fifth trial was only slightly dependent on the response of the previous trial. This indicates that associative learning was not occurring; the children were learning the appropriate response terms, but were unable to connect these terms with specific stimuli.

Context-Learning

The results were similar to those of the previous test. The percentage of correct responses rose from 39% on the first trial to 64% on the second trial and then remained constant over the remaining trials. Of those items correct on the fourth trial, 68% were correct on the fifth trial, compared with the unconditional percentage correct on the fifth trial, 62%. Again, whether or not an item was correct on the fourth trial was not predictive of performance on the fifth trial. This was a recognition test and so the errors can again be attributed to a breakdown in association.

The training series on the consonant-vowel pairs was preceded by pretraining on the vowels, but there was little evidence that the children used this information. A response error could be classified as a vowel error (DE chosen for /de/), a consonant error (CA chosen for /de/), or both (CE chosen for /de/). If the vowel component had been learned, vowel errors should have been rare; a similar argument can be made with respect to consonants. Actually, the even distribution of errors indicates the children treated the items as wholes (32% vowel errors, 34% consonant errors, and 34% both). A similar pattern was found in the transfer data. The percentage of correct answers was 29% (compared with a chance rate of 25%), and the errors were 39% vowel, 36% consonant, and 25% both. The difference during training between variant and invariant items (D vs. C) was negligible: 56% vs. 54% correct, respectively. This finding is also consistent with the thesis that the stimuli were processed as wholes.

In short, there was little evidence of associative learning either of patterns or components after the first trial nor was there any substantial amount of transfer. On the other hand, some children performed quite well on the training series: 7 out of 21 reached a criterion of two perfect trials. Moreover, there was some evidence of differential transfer: those children reaching criterion during training were correct 43% in transfer; those who failed to reach criterion were correct only 20% in transfer.

Letter-Word

Only in this test was there evidence of substantial learning over trials. The results are shown in Fig. 2. Performance on the Letter and Word lists was essentially the same for both the Letter-Word and Word-Letter orders and so the data have been combined over orders. As can be seen from the figure, in the Letter list there was a steady increase over trials in the percentage of correct responses, primarily due to a reduction in omissions. Of those pairs which were correct on the fourth trial, 88% were also correct on the fifth trial, compared with an overall percentage correct on the fifth trial of 66%. In other words, in the Letter list subjects retained previously "learned" items from one trial to the next and added to this

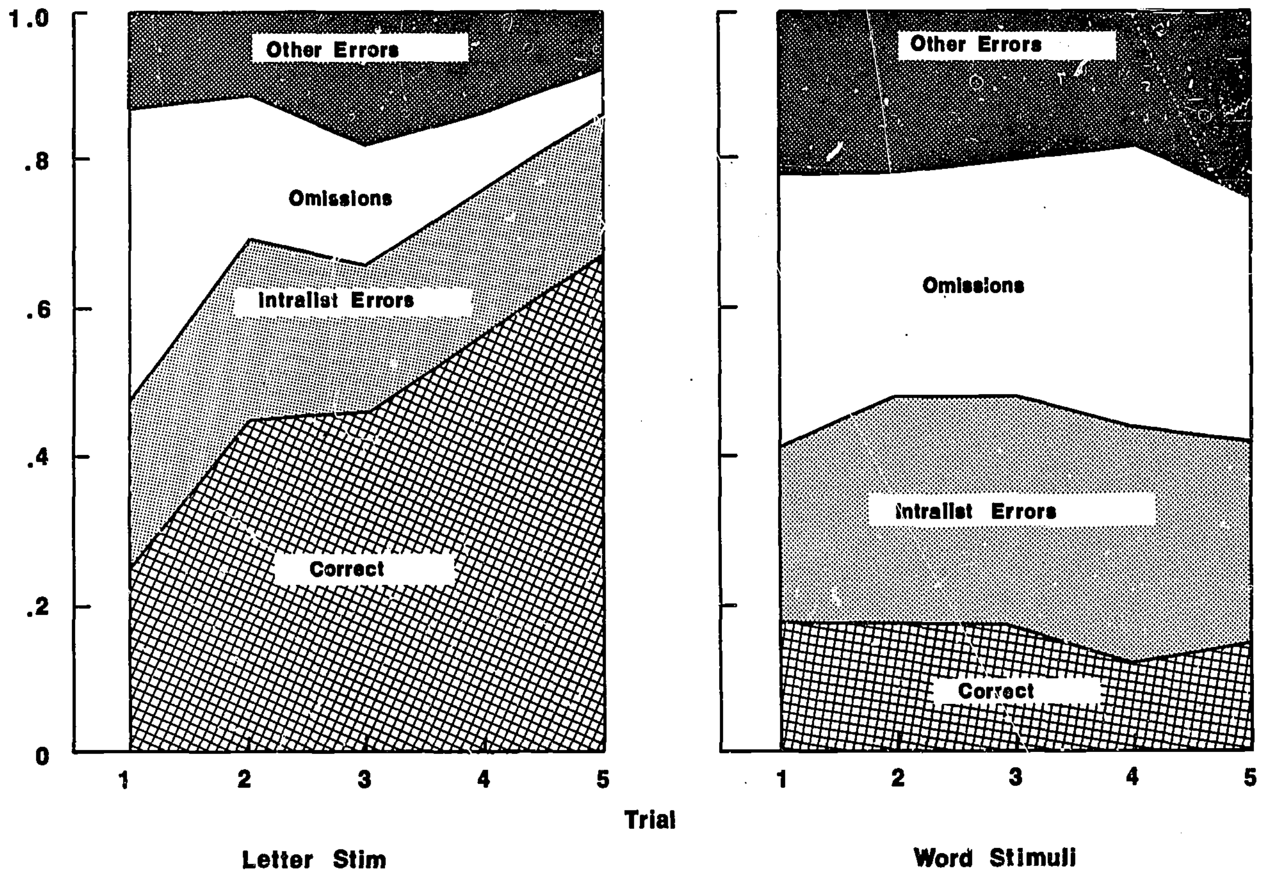

Fig. 2 Letter-Word Test Responses to Letter and Word Lists. Over Trials

store of learned items on each trial. The pattern of results for the Word list, on the other hand, was similar to that previously described for the Alphabet and Context tests. None of the performance measures changed to any noticeable degree over the five trials.

Summary

The kindergartners tested found it extremely difficult to form associations between the unfamiliar abstract stimulus-response pairs used in this portion of the BST package. The only exception was the Letter list of the Letter-Word test. The major distinctions between this latter list and the others were length (three pairs) and the elemental nature of the stimuli and responses (/m/, /n/, and /o/). The effect of list length on paired-associate learning is well known but in the BST package the variation in length was relatively slight (three pairs in the Letter list to six pairs in the Alphabet list). Except for the Letter list, the spoken segment comprised two phonemes in these tasks. Moreover, there was repetition of phonetic elements (e.g., /mo/, /no/, /om/, /on/) producing high intralist similarity, which after the fact we feel may be a critical variable in these particular tests.

In any event, if a child does not "know his ABC's," if he has not learned to identify the letters of the alphabet with their corresponding "names," it is evident that he faces immediately a paired-associate task of some magnitude in beginning reading. Note that the names of the letters of the alphabet are, by and large, difficult in the same sense that our materials were; they are combinations of two (or more) phonemes used in repetitive combinations producing high intralist similarity.

DISCUSSION

Distinctive Feature vs. Pattern Learning

Developmental studies of verbal learning are relatively scarce, given the volume of research in this area on college students. As Flavell and Hill (1969) point out in a recent review, child-learning studies have focused on a small number of specific problems—discrimination, transposition, probability learning, and paired-associate learning—and in none of these areas have the studies revealed anything of much interest relative to development per se. In fact, Flavell and Hill direct their most critical comments toward this lacuna. They call for research which will in "more direct and sensitive ways assess . . . the child's knowledge of classes and relations, his attentional biases, his perceptual skills, and his mnemonic strategies . . . and capacities . . . (p. 44)."

In a particularly revealing experiment, Ann Pick (1965) taught kindergartners to select standard visual symbols from a set of alternatives which included transformations of the standard. For example, if b was the standard, then the set b, d would involve a right-left transformation. [The letters b and d are for illustration only; Gibson's forms were used as stimuli.] After a training criterion was reached, a transfer test was administered. In the Control group, the transfer task included new standards and new transformations (e.g., select w from w, m,). In the Standard group, the same standard set was used but with new transforms (e.g., select b from b, q). Finally, in the Transform group, there was a different standard set, but the transformation remained the same (e.g., select p from p, q).

It was reasoned that if children had learned to choose the standard by reference to an exact template or image in memory, then the Standard group should perform best on the transfer test, since the same templates could be used. If selection was based on transformation of specific features, then the Transform group should do best in transfer. If specific inter-item comparisons were learned, then all three groups should perform equally well on transfer. Pick found that during transfer the Transform group made 40% as many errors as the Control group and the Standard group made 70% as many errors. These results indicate that young children can abstract transformations of distinctive features and use these apart from the specific context in which they were originally embedded.

Transfer of Letter-Sound Learning

There remains the intriguing question of why such transfer occurs under some conditions and not others. In the Letter-Word test in the BST package, there was no evidence of transfer based on combination of elements. Silberman (1964) has recounted his sad experiences in trying to teach kindergartners "reading" matrices with the structure below.

VOWEL-CONSONANT ENDINGS

	-an	-it	-at	-in
f-	Test	Train	Train	Train
r-	Train	Test	Train	Train
s-	Train	Train	Test	Train
m-	Train	Train	Train	Test

In the experimental procedure, the child learned the Train CVC's (e.g., f + at = /fæt/) and then were given a transfer test on the diagonal test items (e.g., f + an = /fæn/). An initial version of the procedure yielded no transfer at all. In the version which finally evolved, the children's attention was directed to the component features (C- and -VC) and then they were taught to amalgamate the elements. A final training segment provided direct practice on transfer to new combinations. After this highly structured program, children performed reasonably well on transfer items (75% correct).

It is apparently unwise to assume that transfer will occur "naturally" or that a process of induction or stimulus generalization guarantees that children will apply information gained in one situation to other similar problems. This may happen, but it may not. In particular, little transfer occurs when the stimulus-response pairs consist of acoustic-phonetic materials. This conclusion is consistent with the earlier discussion of the integrated character of the spoken language.

Other Studies of Letter-Sound Association

A study by Jeffrey and Samuels (1967) bears on "learning the ABC's" and transfer of elements in paired-associate learning involving spoken materials. Jeffrey and Samuels asked whether training on elements or patterns was the more efficient way of teaching reading: is more transfer achieved by teaching words (MO - /mo/, SO - /so/, BA - /be/, BE - /bi/), or letters (M - /m/, S - /s/, A - /e/, E - /i/) where the transfer set consists of words (ME - /mi/, SE - /si/, SA - /se/, MA - /me/)? In the former procedure, the child is trained on integrated word-like units and must induce the letter-sound regularities, whereas in the latter procedure the regularities are taught but the child has less experience in working with integrated units. The kindergarten subjects were first trained to scan letter groups from left to right and to "blend phonically" (e.g., (/m/ - /i/) equals /mi/). Different groups of subjects were given Word training, Letter training, or were trained on an unrelated Control list of cartoon animals paired with proper names. The "letters" used as stimuli were fairly complex visual symbols.

Two results from this study are of interest. First, in reaching a criterion of one perfect trial in training, the Letter group required 13.8 trials; the Word group, 16.9 trials. Compared with the differences in performance over five trials in the Letter and Word lists in BST, this is a small difference. The Jeffrey-Samuels materials were of lower intralist similarity and this may account for the differing results. Second, the number of trials to criterion on the transfer task was 13.5 in the Letter group, 27.2 in the Word group, and 29.3 in the Control group. Thus, training on letter elements substantially facilitated transfer to words in the Jeffrey-Samuels study, although it did not in the Letter-Word task of BST. Several differences in procedure between the two studies might be responsible for the different results. First, in the BST Letter-Word task, original learning was not carried out to a criterion as in the Jeffrey-Samuels study. This is probably not important because subjects who did reach criterion on the Letter-Word test showed no more transfer than noncriterion subjects. Second, in the Jeffrey-Samuels study the child was instructed at the beginning of the transfer task, "You can tell what the word is by making each letter sound. Make the first letter sound, then the next letter sound. Then try to say the word." It would appear than for kindergartners to learn to "read words" in miniature reading tests (a) the child should first be taught individual letter-sound correspondences, (b) it must be made explicit that words are to be read as sound combinations, and (c) instruction in phonic blending must be given. These steps are necessary because young children do not tend naturally to break up reading patterns into elements, nor do they naturally induce correspondences in the process of learning specific examples representing particular correspondences.

VII
VOCABULARY

In one guise or another, a child's knowledge of vocabulary is an important predictor of reading achievement. Vocabulary tests often require children to identify uncommon words, or to try to label indistinct or ambiguous line drawings. Performance on such tests reflects the level of general intellectual ability as well as vocabulary skills.

Three different tests of vocabulary knowledge were included in the BST package. Each child was asked to name common objects displayed in full-color pictures, label line drawings, and to sort the line drawings into the most appropriate categorial arrangements. None of the objects was likely to be completely novel to any child, but the names of some were far less common than the names of others.

METHOD

In Picture-Naming, the child was asked to name each of 22 common objects. The objects were photographed in color and the transparencies were seen in a hand-held viewer In Line-Drawing-Naming, the child was asked to label objects shown as line drawings printed on 3 x 5 cards. The sketches included some detail and shading. There were five objects in each of six categories and six objects in four categories. The categories were animals, food, toys, vehicles, articles of clothing, furniture, dinnerware, parts of the body, insects, and round things. Within these categories, items were selected on the basis of distinctiveness and familiarity.

The line drawing cards were also used in the Sorting test, which consisted of three parts: (a) two categories without exemplars, (b) four categories with exemplars, and (c) four categories without exemplars. In each of these subtests, the child was given a deck of cards containing five cards from each category and asked to sort them into groups that "belong together." A formboard was used to force the children to sort into the desired number of categories. In (b), four exemplar cards were placed at the left of each row of the formboard, and the child was instructed to sort the deck into four groups using the exemplars as a basis.

RESULTS

The error rate on Picture-Naming was low: 4.1% (Msn) and 3.7% (Bel). Of the total of 39 errors, 17 were intraclass errors of one sort or another (e.g., rose for flowers, pie for cake), 6 were omissions, and the remaining 16 were "other" (e.g., necktie for hat, bread for apple, cookies for pennies).[6] The errors were widely distributed over subjects—there were 17, 16, 7, and 3 children who made 0, 1, 2, or 3 errors respectively.

The error rate on Line-Drawing-Naming was higher than on Picture-Naming, largely because the test contained more unfamiliar items. On the eight objects common to the two tests (apple, cake, chair, car, cat, dog, horse, shoe) there were actually more errors on Picture-Naming than on Line-Drawing-Naming (12 versus 4). Overall, the error rate was 15.7% (Msn) and 20.6% (Bel) on Line-Drawing-Naming. Although the rate was slightly higher in Bel than Msn, the pattern of errors was the

---

[6] The scoring scheme used to classify these responses is admittedly arbitrary in some instances. In general, our classification was conservative; by a more liberal criterion, only two or three of the responses were not related semantically or perceptually to the stimulus. For example, one might argue that necktie for hat and cookies for pennies bear some relation to the stimulus, even though we classed them as "other."

same in both samples, and so analysis will be based on the combined data.

Failures to respond comprised 15% of the errors. One quarter of the errors were accounted for by five popular substitutions; more than a third of the children said wheel for tire, dresser or drawer for desk, baby for doll, windmill for pinwheel, or crib for bed. A closely related error was a coordinate or intraclass confusion, such as penny or nickel for half dollar, spider for bee, or goat for cow. (The distinction between popular and coordinate confusions was based on frequency of intrusion and definitional equivalence.) These accounted for an additional 34% of the errors. Another 12% were errors of superordination or overgeneralization, such as money for half dollar, meat for hot dog, or face for mouth. The remaining 14% of the errors were classified "other." Most of these latter errors were the result of overdifferentiation of the stimulus (e.g., bird to half dollar), a few others were incomprehensible (e.g., leaf for ant, "boodle allen" for deer).

Seven words (half dollar, mouth, bee, ant, grasshopper, spider, and pitcher) produced a large proportion of errors marked by considerable variation (e.g., pitcher was variously described as coffee pot, coffee cup, cup, can, kettle, bottle, cream bowl, etc., ant as grasshopper, spider, bug, fly, flea, leaf, dirt, etc.). More than a third of the children gave deviant responses to these cards. On 28 of the remaining 38 stimuli, four children or less made responses classed as errors. Thus, the distribution of errors over the stimuli tended to be bimodal: stimuli produced many or few errors. The distribution of errors over subjects tended to be unimodal, contrary to the bimodal subject distributions found in other tests. In Line-Drawing-Naming, errors ranged from 1 to 15, with 60% of the error scores between 5 and 10. The split-half reliabilities were .53 (Msn) and .80 (Bel).

Sorting

This test was given to the Msn sample only. One performance measure was the number of times the most frequent category in each row was represented. This measure disregarded the cue cards in the four-category-with-exemplar subtest. The guessing rate was 50% for the two-category subtest and 40% for the four-category subtests. The observed percentages of correct responses were 70%, 64%, and 63% for the two-category, four-category-with-exemplar, and four-category-without-exemplars subtests, respectively. The percentage correct on the second test dropped to 56% when correct responses were defined as those sortings which matched the exemplar.

There were substantial individual differences on the sorting task, showing up as a markedly bimodal distribution. About half of the children (10 out of 21) made scores between 37 and 48 correct out of 50 possible, which is less than one and a half mistakes per category. The other 11 children made scores of 27 to 22 correct, or about two and a half mistakes per category.

The distribution of sorting errors over the stimulus categories, on the other hand, was uniform within each of the two four-category sets used in the test, although there was a sizable difference between the sets. The relevant data are presented in Table 8. The frequencies along the main diagonals are correct sorts, the number of times that two items belonging to the same normative category were placed in the same stack by a child. The off-diagonal frequencies are category confusion errors; thus, in set 1 there were 63 occasions in which an item from the kitchen category was placed in the same stack as an item from the clothes category. Also shown in the table are the number of labeling errors for items in each category by children in the Msn and Bel examples.

Several aspects of the data should be noted. First, within each set there was little evidence of differences in the number of correct sorts for any category. There was a difference between sets, which could have resulted from specific category difficulty, or differential intracategory confusability or both. For example, the toy category had the lowest number of correct sorts and was also more frequently confused with other categories, particularly body parts and furniture. This result could reflect a poorly developed concept of toy, or the existence of specific confusions between toy-body parts and toy-furniture; the data from this study do not permit us to decide between these two interpretations. Second, the distribution of confusion errors is reasonably uniform. As just noted, toy-body parts and toy-furniture confusions are greater in number than other combinations, but even these maximum deviations from a uniform distribution of confusion errors are relatively slight. Third, there is no discernible relation between sorting performance and labeling performance. In Set 1, there were 21 labeling errors (disregarding "populars") for kitchen items and 2 labeling errors in the vehicle category; the number of correct sorts in these two categories was virtually identical. Similarly, in Set 2 there was a large difference in labeling errors for the toys and furniture categories, yet only a slight difference in sorting

Table 8

Correct Sorts and Intercategory Sorting Confusions in Four-Category Sets

Set 1

Kitchen (21, 28)	Clothes (9, 12)	Animals (8, 18)	Vehicles (2, 6)	Categories
112	63	70	63	Kitchen
	111	70	69	Clothes
		106	64	Animals
			117	Vehicles
Toys	64			
Food	79	84		
Body Parts	101	77	93	
Furniture	111	90	59	78
Categories	Toys (15, 33)	Food (5, 4)	Body Parts (8, 20)	Furniture (1, 4)

Set 2

NOTE—Numbers in parentheses are labeling errors in Msn and Bel samples.

errors. In short, the coherence of a category was not related to the availability of labels for items in the category for the reasonably familiar and concrete items used in this test. The pattern of labeling errors was similar in the Msn and Bel samples; the Spearman rank-order correlation between labeling errors per category for the two samples was .85.

DISCUSSION

Vocabulary Size

Under the conditions used in the BST package, the children tested generally gave appropriate names for the common objects in the test. When errors occurred, only rarely were these omissions; most commonly, children produced words that were "close" to the normative label. To be sure, the test words were concrete objects presented in isolation. The effects of variation in material (abstract vs. concrete, verbs and adjective forms vs. nouns) and test procedure remain to be fully explored.

Prior to the advent of transformational grammars, estimation of vocabulary size was one of the most popular topics for psycholinguistic investigation (Irwin, 1960; Templin, 1957; McCarthy, 1954). Performance on various types of vocabulary tests were shown to predict intelligence and general school achievement, including reading acquisition. This last relation might be considered unusual in light of the severe limitations on vocabulary size and content in most reading series. Several studies have shown vocabulary size in kindergartners to be substantially greater than the number of different words in most reading series.

It seemed to us that a vocabulary test which used sufficiently direct means of eliciting the names of common objects would not distinguish between good and poor readers. Our assumption was that virtually all children would know the most common words in the language, though they might vary in the extent to which they were familiar with less common words.[7] There is some evidence from BST intertest correlations to be discussed later that this assumption is incorrect, that even in identification of common words the quality or character of a child's label may be related to other cognitive skills which we have assumed to be important in beginning reading.

Vocabulary Quality

A variety of techniques have been used to measure vocabulary: samples of free speech, elicitation of words from pictures of line drawings, recognition tests in which the child hears a word and is asked to pick out the most appropriate picture in a set, asking the child to define a word pronounced by the experimenter, free association tests in which the child is asked to pronounce as many words as he can think of in a fixed period of time, or word association tests in which the child is asked to say the first word that comes to his mind when the tester pronounces a stimulus word. The "vocabularies" being measured by these various techniques might be expected to be quite different.

Templin (1957) distinguished between <u>use</u> and <u>recognition</u> vocabulary. She measured the former by the Seashore-Eckerson Vocabulary Test in which the child is asked to name a picture, and the latter by the Ammons Picture Recognition Test in which the child is shown several pictures and asked to point to the picture corresponding to a word spoken by the tester. A group of 3- to 5-year olds received the Ammons test, and a group of 6- to 8-year olds the Seashore-Eckerson. Each age group included children from upper and lower socioeconomic levels. There was no substantial difference in performance on the Ammons Test between children from upper and lower socioeconomic levels in the age range from 3 to 5. However, 6- to 8-year olds from a lower socioeconomic level performed substantially more poorly on the Seashore-Eckerson Test than an upper socioeconomic group. Although the type of test was confounded with age, the data suggest that the nature of vocabulary test may be an important variable in its own right, again pointing out the necessity of convergent tests.

There are many methodological problems in testing children's vocabulary which have not been systematically examined. We know of no good techniques for testing anything other than concrete nouns, concrete adjectives, and action verbs in young children. The method of elicitation and the choice of whether words are to be tested in a sentential context or in isolation probably affect performance. Finally, the characteristics of the tester are important, especially in testing different social and cultural groups. For example, there is evidence that the speech of Negro children from urban ghettos depends on whether or not they are tested by interviewers of their own race, who may (or may not) speak a dialect closer to the child's (Entwisle & Greenberger, 1968; Labov, 1967).

Although a large amount of research had been completed through 1950 on the development of vocabulary size, techniques of analysis were so unrefined that little can be said about the actual process of development, except that vocabulary size increased with age. Most investigators and reviewers have raised the problem of what it means to "know" a word. If the developing child simply has an increasing number of words at his disposal, that is one thing, however, it has been frequently suggested that not just the number of words available, but the way in which a child is able to use those words changes with age. McCarthy (1954) mentions work by Feifel and Lorge (1950) and by Gerstein (1949) on the form of definitional responses used by children from 6 to 14 years of age in the Stanford-Binet and Wechsler-Bellevue Vocabulary items. "The younger children more often employed the use, description, illustration, and demonstration types of definition as well as inferior explanation and repetition responses. Older children significantly more often used synonym and explanatory responses (p. 530)." Unfortunately, when the definitional type of vocabulary test is used, one cannot tell whether or not the observed behavior is an accurate representation of the language system or whether one is primarily testing the productive system as opposed to the receptive or comprehension system. In other words, younger and older children may understand words equally well, but older children are simply better able to explain their understanding. Communications research by Krauss and Glucksberg (1969a, 1969b) supports the latter interpretation.

Russell and Saadeh (1962) also presented some interesting results on vocabulary quality. They administered a four-alternative-definition recognition test to third, sixth, and ninth grade subjects. The four answers provided as

---

[7] Children for whom English is a second language, or those from extremely deprived backgrounds, might prove exceptions to this generalization.

alternatives were classed as functional, concrete, abstract, and incorrect. [For example, COUNT: to find the number of things in a group (functional), to find how many pennies are in your pocket (concrete), to say numbers in order (abstract), and to tell numbers one after another (incorrect). The validity of the results in this kind of experiment depends greatly on the quality of the alternatives, of course.] The data showed a clear shift from concrete to abstract responses between the third and sixth graders, and this was taken as evidence of improved comprehension.

Word Associations

Another lead to vocabulary functioning comes from the literature on word associations. A number of investigators (Entwisle, 1966; Palermo & Jenkins, 1964; Ervin, 1961; Brown & Berko, 1960) have found a shift from syntagmatic to paradigmatic word associations from about 5 to 7 years of age. [Paradigmatic responses are those in which the form class of the response matches that of the stimulus, such as black-white or boy-girl; syntagmatic responses are those of a different form class which could precede or follow the stimulus word in a sentence, such as black-dog or boy-hit.] This shift can be interpreted as a change in the semantic structure of the child's language systems, or a change in production ability. Data on this point was obtained by Ervin (1961) who tested children with both word association and two-choice recognition procedures. In the latter test, children selected the more appropriate of two associations for a given stimulus. Certain stimuli were particularly sensitive to the syntagmatic-paradigmatic shift, e.g., ball: bat or play. In kindergarten and first grade children, the syntagmatic response (play) was selected 67% of the time, while sixth graders chose the paradigmatic response (bat) 57% of the time. Thus, even when production ability is minimized, the shift is observed. The corresponding percentages in the productive word-association test were 71% and 61%.

The thoughtful work of Entwisle (1966) on word associations in young children is worth noting, especially since it spans the age range of most interest to us. In her concluding remarks, Entwisle is quite specific as to what word associations are not: they are not simple stimulus-response linkages; they are not elements of serial chains taken from language samples; they are not representative of more than limited portion of the child's language; and they do not directly reflect the semantic space of the child. On the positive side, Entwisle makes five major points in her summary:

(a) The major shift from syntagmatic to paradigmatic associations around the first grade was confirmed but found to be much more complex than previously realized.
(b) The age at which the shift occurred depended on form class and word frequency. For adjectives with antonymic contrasts, such as bright-dark, the shift occurs earlier than for those without such contrasts, such as thirsty-?, or for verbs which as a class do not provide many natural contrasts.
(c) There are asymmetries in the form-class relations; adverbs produce adjectives, but not vice versa.
(d) The syntagmatic responses of children differ qualitatively from those of adults; e.g., noun-verb sequences are used by children, whereas adults are more prone to produce adjective-noun pairs.
(e) Children give a greater variety of associations than do adults.

Entwisle, following a line of argument also developed by McNeill (1966), feels that these results, particularly (a), (b), (c), shed light on the development of the semantic system;" . . . [they] tend to reveal the formation of word classes or concepts and so they forecast the individual's potential ability to emit different combinations of words from those he has heard" (Entwisle, 1966, p. 7). She rejects the idea that in associating to a stimulus word the subject finds the word in a mental dictionary and pulls off the topmost entry in an associative list or response hierarchy. Instead, it is assumed that a word in memory consists of a feature list (Katz & Fodor, 1963) containing both semantic and syntactic features. Hence, dog might be described as animal, domestic, . . . common noun, regular plural, ... With increasing age, this system of features becomes more fully developed and better organized, so that for an adjective such as hard, a subject has immediate associative access to the antonym soft because it differs by a single feature. Prior to the full elaboration of this system, there may be confusions among similar words, because the entries in the feature list are not well enough established to permit rapid discrimination of close associates.

An aspect of Entwisle's data of particular interest to us which fits naturally into a feature-list theory is the presence of many acoustic-phonetic or "clang" associates,

GPO 818-669-4

particularly in the responses of younger children. If the phonetic description of a word is stored in memory along with the semantic and syntactic features (see Brown & McNeill, 1966), one would expect phonetic associates to occur, particularly if other associations are not readily available. Such associative rhymes were produced frequently by urban Negro children (Entwisle & Greenberger, 1968), by Entwisle's (1966) kindergartners, by Palermo and Jenkins' (1964) sample of fifth grade through college students, in Ervin's (1961) Kindergarten to sixth grade sample, and in a free-association study in which college students were asked to produce three-letter words as fast as possible (D. Nelson, personal communication; a sample might be man, tan, ran, fan, fad, bad, sad, etc.). In Entwisle's (1966) study, for example, the percentages of rhyming responses to add (a high frequency, familiar word) by kindergarteners, first, second, and third graders were 13%, 7.5%, 3%, and 1%, respectively. For bitter (low-frequency), the corresponding percentages were 20%, 18%, 11%, and 3%. Ervin obtained similar shifts in the rate of phonetic associates from kindergarten to sixth grade. As she remarked, these age-related trends parallel the shift from phonological to semantic conditioning reported by the Russians (Razran, 1939) and by Riess (1946). In the data of Palermo and Jenkins (1964), words such as the, at, and now, which are frequent but have relatively little semantic content, evoked phonetic associates occasionally even in college students (percentages of rhyming associates produced by college sophomores were 3%, 7%, and 4%, respectively, for the stimulus words given above).

Entwisle and Greenberger (1968), testing Negro children from the Inner City of Baltimore, found that the primary associations by first grade Negro children were frequently rhyming responses, such as bad, fad, or mad to the stimulus word add, or mean and bean as responses to clean. These responses usually bore no semantic relationship to the stimulus word, and in many instances the responses were nonsense, such as fird to bird. In the case of urban Negro children, in 55 of the 96 stimulus words in the Entwisle-Greenberger list, one or more of the three most common associates was a rhyming word. The corresponding count for urban white children was 18 out of 96.

The word association task provides a fruitful method for examining semantic aspects of the language system. The frequency of phonetic associations prior to about 6 years of age is especially interesting. If it is true that analysis of the acoustic properties of spoken language does constitute a major hurdle to the learning of the decoding portion of reading, and if the preschool child is more inclined than the kindergartner or first grader to focus on the phonetic composition of words, the teaching of reading (at least that portion concerned with decoding) might be more efficiently introduced at age four or five.

VIII
INTERTEST CORRELATIONS

RESULTS

One of our aims in the BST package was to develop clusters of tests which converged on specific skill areas—visual, acoustic-phonetic, learning, and vocabulary. Since performance was poor on most of the acoustic-phonetic and learning tests, few intertest correlations in those areas will be reported. Intertest correlations for the tests with satisfactory performance levels are presented in Table 9, grouped according to skill areas. Alphabet-Production and Alphabet-Recognition have been placed with the visual matching tasks since the latter used real letters as stimuli. Three other variables of general interest—age, Metropolitan Readiness percentile, and word memory span—have been included in the table.

The following conclusions summarize the matrix of correlations. Clusters of tests pertinent to these conclusions are enclosed in block form in Table 9.

Visual Tasks and Alphabet Knowledge

The visual matching tests in the BST package did hang together quite well. Of the seven intercorrelations involving Tests 2, 3, and 4, six were significant at the .01 level, the other at the .05 level.

There was some relation between knowledge of the alphabet and visual matching. Of the twelve intercorrelations in this set, for six $p < .01$, for one $p < .05$, and the other five ranged from .20 to .36. Oddity-Selection and Memory-Matching were least related to alphabet knowledge. (Alphabet-Production and Recognition were significantly related, it will be noted.) The relation of performance within the visual matching tests was closer than the relation between visual matching and alphabet knowledge, which indicates that the visual matching tests tapped a component skill other than general competence or age. This conclusion would be on firmer footing if visual matching performance with non-alphabet stimuli such as Gibson forms were tested.

Vocabulary

The predicted relation among the vocabulary tests in the BST package was not found. The absence of a correlation between Picture-Naming and Line-Drawing may reflect the low error rate on the former test. The fact that labeling and sorting were unrelated was noted in the preceding discussion on the Sorting test.

Acoustic-Phonetic Tasks

Segmentation and Rhyme-Production showed inconsistent patterns of significant intertest correlation in the two S groups when only correct segmentation responses on Trial 5 and transfer were scored. When phonologically related responses were combined with correct (as in Tables 6, 7, and 9), the strongest relation was between Segmentation and Rhyme-Production.

Other BST Correlations

There were 77 other correlations among BST tests (age excluded) for which no relation was predicted; of these, eight were significant at the .01 level, five at the .05 level, and the remaining were not statistically significant. Of the 29 BST test pairs for which both Msn and Bel correlations could be computed, only one showed significant correlations for both groups. In other words, most of the tests were reasonably independent of one another. In particular, word-memory span bore no significant

Table 9

Intercorrelations Among Selected Tests from the BST Package

	2	3	4	5	6	7	8	9	10	11	12	13	14	m	s
1. Age (Mo.)	36 -01	40* -20	22 --	35 --	68** -08	27 -01	16 35	23 04	52** --	24 06	28 14	-16 63**	34 24	64 69	4 3
2. Matching Letter groups		79** 72**	58** --	40* --	50** 41*	58** 67**	36 34	27 35	01 --	01 54**	17 13	08 09	36 67**	63 68	11 12
3. Matching Retest			53** --	52** --	50** 34	63** 75**	60** 46*	48* 34	08 --	09 61**	13 33	12 18	51** 68**	49 42	25 32
4. Oddity Selection				63** --	24 --	36 --	33 --	22 --	25 --	25 --	-19 --	20 --	22 --	26 --	23 --
5. Memory Matching					20 --	35 --	21 --	29 --	27 --	20 --	07 --	05 --	45* --	43 --	19 --
6. Alphabet Production						61** 45*	27 05	35 04	46* --	26 22	48* 35	04 -28	42* 21	27 33	27 31
7. Alphabet Recognition							34 36	59** 30	15 --	15 52**	29 35	17 03	46* 57**	23 18	31 27
8. Picture Naming								18 08	11 --	04 29	13 27	17 -01	10 48*	95 95	5 5
9. Line-Drawing Naming									32 --	29 59**	15 54**	15 32	67** 36	82 78	7 7
10. Sorting										42* --	24 --	08 --	46* --	61 --	16 --
11. Segmentation											54** 71**	31 19	36 58**	66 59	6 8
12. Rhyming Production												03 04	25 24	39 37	40 43
13. Word-Memory Span													24 14	8 8	1 2
14. Metropolitan (Percentile)														62 41	20 27

NOTE—Upper entry in each cell is Msn sample, lower entry is Bel sample. N = 21 for each sample, one S in Bel not included because test 14 was not available. Decimals omitted. m is per cent correct except for tests 1, 13 and 14; S is the standard deviation.

* $r > .37$, $p < .05$, one-tailed test. ** $r > .50$, $p < .01$, one-tailed test.

relation to performance on any of the other tests. (Many of the correlations are between .20 and .35; it may be a general law of behavioral testing that any two variables will jointly account for 5% to 10% of the total variance.) Age did not appear to be consistently related to performance on most tests.

Metropolitan Tests

For comparison, the intercorrelations among the six subtests and percentile rank of the Metropolitan Readiness Tests are shown in Table 10. Although each subtest was designed to tap an independent skill, 14 of the 30 correlations were significant at the .01 level and 5 at the .05 level. The remaining 10 were not statistically significant. Of the 19 correlations of Metropolitan Readiness percentile with BST tests, 6 were significant at the .01 level, 6 at the .05 level, and 7 were not statistically reliable. In short, although the subtests of the Metropolitan were not independent, the test served well as a general predictor of performance in a variety of cognitive skills.

DISCUSSION

Cronbach (1967) has recently raised the possibility of a marriage between differential and experimental psychology. Several contributors to the volume (Gagné, 1967) containing Cronbach's paper take note that these two areas have enjoyed an engagement of long duration but there is no evidence that the affair has ever been consummated. The failure arises from the difficulty of simultaneously manipulating stimulus variables, training variables, task variables, and subject variables. Without pretending that our effort represents more than a limited step in this direction, some of the results are promising. First, without benefit of factor analysis, the BST data in Table 8 do hang together in _a priori_ clusters when they hang together at all. The majority of the other intercorrelations are negligible in one, the other, or both samples. Second, distributions of subject or stimulus scores are bimodal in a number of tests, which permits partitioning of subjects or stimuli in a relatively simple pass-fail fashion. The questions then follow: (a) why do some subjects

Table 10

Intercorrelations Among Subtests of Metropolitan Readiness Tests

	2	3	4	5	6	7	_m_	_s_
1. Word Meaning	60** 56**	06 69**	30 52**	48** 76**	16 51**	68** 70**	63 52	17 16
2. Listening		05 46*	26 43*	35 53*	20 53**	62** 64**	64 59	15 13
3. Matching			05 62**	40* 63**	47* 69**	53** 85**	69 49	18 25
4. Alphabet				50** 71**	-04 58**	56** 84**	85 40	16 27
5. Numbers					26 56**	82** 83**	45 38	12 17
6. Copying						55** 79**	42 60	19 27
7. Percentile							62 41	20 27

NOTE—Upper entry in each cell is _Msn_ sample, lower entry is _Bel_ sample. N = 21 for each sample, one _S_ in _Bel_ not included because test 14 was not available. Decimals omitted. _m_ is per cent correct; _S_ is the standard deviation.

* $r > .37$, $p < .05$, one-tailed test.

** $r > .50$, $p < .01$, one-tailed test.

perform well and others poorly and (b) why is the one group of stimuli satisfactorily handled and the other not? In both instances, the answers will take the form of experimental manipulations—training procedures or variations in materials which produce satisfactory performance. It is important to note that feedback under these conditions is immediate—you quickly find out when a testing or training procedure doesn't work.

IX
SUMMARY

Even though the BST package is in a preliminary form, and despite the limited number of subjects, certain conclusions are worth noting. First, if skill components are narrowly defined, they appear remarkably independent. This independence is not the result of suppressing variation among individuals; from the bimodal distribution of Ss in Rhyme-Production one might well expect high correlations with other tasks, and yet these are not found except for the Segmentation test.

With regard to specific skill areas, it appears that visual perception skills contribute only minimally to matching tests. Few errors are made in the matching of single letters, where perceptual problems should play the primary role. Instead, the problems which arise seem to be of a cognitive nature, such as order and memory for forms. The data suggest that sound matching, segmentation, and association of sounds and symbols are poorly developed skills in most kindergartners. However, testing procedures may be at fault —a possibility which must be evaluated more fully.

Finally, two comments can be made about the relation of the BST data to reading achievement. First, correlational data on this relationship will be available in a year and the relation can be determined empirically. Second, inquiries about the relation of specific component skills to reading achievement may be irrelevant to our purpose because most achievement tests are not particularly sensitive to the decoding process on which BST package focuses.

REFERENCES

Alpern, G. D. The failure of nursery school enrichment program for culturally disadvantaged. American Journal of Orthopsychiatry, 1966, 36, 244-245.

Battig, W. F. Procedural problems in paired-associate learning research. Psychonomic Monograph Supplement, 1965, 1, No. 1 (Whole).

Benton, A. L. Right-left discrimination and finger localization. New York: Harper, 1959.

Bereiter, C., & Engelmann, S. Teaching disadvantaged children in the preschool. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Bond, G. L. The auditory and speech characteristics of poor readers. Teachers College Contributions to Education, No. 657. New York: Columbia University, 1935.

Bond, G. L., & Dykstra, R. The cooperative research program in first-grade reading instruction. Reading Research Quarterly, 1967, 2 (complete).

Brottman, M. A. Language remediation for the disadvantaged preschool child. Monographs of Society for Research in Child Development, 1968, 33, (Whole No. 124).

Brown, R., & Berko, J. Word association and the development of grammar. Child Development, 1960, 31, 1-14.

Brown, R., & McNeill, D. The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 325-337.

Bruce, D. J. The analysis of word sounds by young children. British Journal of Educational Psychology, 1964, 34, 158-170.

Chall, J. Learning to read: the great debate. New York: McGraw-Hill, 1967.

Cronbach, L. J. How can instruction be adapted to individual differences? In R. M. Gagne (Ed.), Learning and Individual Differences. Columbus, Ohio: Charles Merrill, 1967.

Davidson, H. A study of reversals in young children. Journal of General Psychology, 1934, 45, 452-465.

Deutsch, C. P. Auditory discrimination and learning: social factors. In M. Deutsch (Ed.), The Disadvantaged Child. New York: Basic Books, 1967.

Dunn-Rânkin, P. The similarity of lower-case letters of the English alphabet. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 990-995.

Elkonin, D. B. The psychology of mastering the elements of reading. In B. & J. Simon (Eds.), Educational Psychology in the U.S.S.R. London: Routledge & Kegan Paul, 1963.

Entwisle, D. R. Word associations of young children. Baltimore, Md.: Johns Hopkins Press, 1966.

Entwisle, D. R., & Greenberger, E. Differences in the language of Negro and white grade-school children. Technical Report 19, Center for study of social organization of schools. Baltimore, Md.: Johns Hopkins University, 1968.

Ervin, S. M. Change with age in the verbal determinants of word-association. American Journal of Psychology, 1961, 74, 361-372.

Feifel, H., & Lorge, I. Qualitative differences in the vocabulary responses of children. Journal of Educational Psychology, 1950, 41, 1-18.

Fellows, B. J. The discriminative process and development. New York: Pergamon, 1968.

Flavell, J. H., & Hill, J. P. Developmental Psychology. Annual Review of Psychology, 1969, 20, 1-56.

Gagné, R. M. (Ed.). Learning and individual differences. Columbus, Ohio: Charles Merrill, 1967.

Gerstein, R. A. A suggested method for analyzing and extending the use of Bellevue-Wechsler vocabulary responses. Journal of Consulting Psychology, 1949, 13, 366-370.

Gibson, E. J. Learning to read. Science, 1965, 148, 1066-1072.
Gibson, E. J., Gibson, J. J., Pick, A. D., & Osser, H. A developmental study of the discrimination of letter-like forms. Journal of Comparative and Physiological Psychology, 1962, 55, 897-906.
Gibson, E. J., & Yonas, A. A developmental study of visual search behavior. Perception and Psychophysics, 1966, 1, 169-171.
Haber, R. N. Perception and thought: An information-processing analysis. J. F. Voss (Ed.), Approaches to Thought. Columbus, Ohio: Charles Merrill, 1969.
Harris, L. Discrimination of left-right directionality and development of the logic of relations. Paper presented to Society for Research in Child Development, Santa Monica, California, March, 1969.
Hildreth, G., Griffiths, N. L., & McGaurran, M. E. Metropolitan readiness test. New York: Harcourt, Brace and World, 1964.
Hill, M. B. A study of the process of word discrimination in individuals beginning to read. Journal of Educational Research, 1936, 29, 487-500.
Howard, I. P., & Templeton, W. D. Human spatial orientation. New York: Wiley, 1966.
Huttenlocher, J. Children's ability to order and orient objects. Child Development, 1967(a), 38, 1169-1176.
Huttenlocher, J. Discrimination of figure orientation: effects of relative position. Journal of Comparative and Physiological Psychology, 1967(b), 63, 359-361.
Irwin, O. C. Language and communication. In P. H. Mussen (Ed.), Handbook of research methods in child development. New York: Wiley, 1960.
Jeffrey, W. E. Variables in early discrimination learning: I. Motor responses in the training of a left-right discrimination. Child Development, 1958, 29, 269-275.
Jeffrey, W. E., & Samuels, S. J. Effect of method of reading training on initial learning and transfer. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 354-358.
Kamil, M. L., & Rudegeair, R. E. Assessment of phonological discrimination in children. Paper presented at Midwestern Psychological Association, Chicago, May, 1969.
Katz, J. J., & Fodor, J. A. The structure of a semantic theory. Language, 1963, 39, 170-210.
Krauss, R. M., & Glucksberg, S. The development of communication: Competence as a function of age. Child Development, 1969(a), 40, 255-266.
Krauss, R. M., & Glucksberg, S. Some characteristics of children's messages. Paper presented at Society for Research in Child Development, Santa Monica, California, 1969(b).
Labov, W. Some sources of reading problems for Negro speakers of nonstandard English. In A. Frazier (Ed.), New directions in elementary English. Champaign, Ill.: National Council of Teachers of English, 1967.
McCarthy, D. Language development in children. In L. Carmichael (Ed.), Manual of child psychology. New York: Wiley, 1954.
McNeill, J. D., & Stone, J. Note on teaching children to hear separate sounds in spoken words. Journal of Educational Psychology, 1965, 56, 13-15.
McNeill, D. A study of word association. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 548-557.
Monroe, M. Children who cannot read. Chicago: University of Chicago Press, 1932.
Muehl, S., & Kremenak, S. Ability to match information within and between auditory and visual sense modalities and subsequent reading achievement. Journal of Educational Psychology, 1966, 57, 230-239.
Neisser, U. Cognitive psychology. New York: Appleton-Century-Croft, 1967.
Palermo, D. S., & Jenkins, J. J. Word association norms. Minneapolis, Minn.: University of Minnesota Press, 1964.
Pick, A. D. Improvement of visual and tactual form discrimination. Journal of Experimental Psychology, 1965, 69, 331-339.
Pufall, P. B., & Furth, H. G. Recognition and learning of visual sequences in young children. Child Development, 1966, 37, 827-836.
Razran, G. H. S. Semantic, syntactic and phonetographic generalization of verbal conditioning. Psychological Bulletin, 1939, 36, p. 578.
Riess, B. F. Genetic changes in semantic conditioning. Journal of Experimental Psychology, 1946, 36, 143-152.
Russell, D. H., & Saadeh, I. Q. Qualitative levels in children's vocabularies. Journal of Educational Psychology, 1962, 4, 170-174.
Silberman, H. F. Experimental analysis of a beginning reading skill. Programed Instruction, 1964, 3, 4-8.
Skeel, M. H., Calfee, R. C., & Venezky, R. L. Perceptual confusions among fricatives in preschool children. Technical Report No. 73, Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, February, 1969.

00047

GPO 818-069-3

Stern, C. Evaluation language curricula for preschool children. In M. A. Brottman (Ed.), Language remediation for the disadvantaged preschool school, Monographs of Society for Research in Child Development, 1968, 33, (Whole No. 124).

Strang, H. R. The effects of letter directionality cueing and response mode upon the acquisition of letter reversal discriminations in four-year-old children. Unpublished Ph.D., University of Kansas, 1967.

Templin, M. C. Certain language skills in children. Minneapolis: University of Minnesota Press, 1957.

Tinker, M. A. Recent studies of eye movements in reading. Psychological Bulletin, 1958, 55, 215-231.

Uhr, L. (Ed.), Pattern recognition. New York; Wiley, 1966.

Venezky, R. L., Calfee, R. C., & Chapman, R. Skills required for learning to read: A preliminary analysis. Education, 1969, 89, 298-302.

Vernon, M. D. Backwardness in reading: A study of its nature and origin. Cambridge, England: University Press, 1960.

GPO 818-669-2