DOCUMENT RESUME

ED 273 934 CS 008 547

AUTHOR	Lesgold, Alan; And Others
TITLE	Learning to Read: A Longitudinal Study of Word Skill Development in Two Curricula. 1986/4.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
REPORT NO	LRDC-1986/4
PUB DATE	85
NOTE	35p.; Reprint from: Mackinnon, G. E., Ed.; and Waller, T. G., Ed. Reading Research: Advances in Theory and Practice. Academic Press, Vol. 4, 1985, p107-138.
PUB TYPE	Reports - Research/Technical (143) -- Journal Articles (080)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Basal Reading; Beginning Reading; *Decoding (Reading); Elementary Education; Individualized Reading; Longitudinal Studies; Oral Reading; *Reading Comprehension; Reading Improvement; *Reading Instruction; Reading Rate; *Reading Research; Sight Method; *Teaching Methods; *Word Recognition

ABSTRACT

A study examined how word recognition automaticity develops and its relationship to the acquisition of comprehension skill. Two different methods for teaching reading were used: (1) a global method using the Houghton Mifflin basal reading program, and (2) a code method using the New Reading System, which emphasizes word decoding skills along with comprehension skills. Core data were collected with children over a three-year period from first through third grade. Measures included reaction times for oral reading of individual words and for judgments of word meanings, comprehension of sentences and passages, coding of errors for sensitivity to context and fidelity to the phonemic code, and subtest scores from achievement tests. Results indicated a relationship between word recognition efficiency early in learning and reading comprehension performance later on. The two different teaching approaches resulted in substantial qualitative differences in acquisition of word recognition ability. These differences in oral reading disappeared by the end of third grade. The findings do not provide a basis for choosing between code and global methods, but they do suggest that neither method sufficiently develops word recognition efficiency. Eight figures and four tables of data supplement the text. (SRT)

LEARNING RESEARCH AND DEVELOPMENT CENTER

LEARNING TO READ: A LONGITUDINAL STUDY OF WORD SKILL DEVELOPMENT IN TWO CURRICULA

1986/4

ALAN LESGOLD, LAUREN B. RESNICK, AND KATHLEEN HAMMOND

University of Pittsburgh

LEARNING TO READ: A LONGITUDINAL STUDY OF WORD SKILL DEVELOPMENT IN TWO CURRICULA

Alan Lesgold, Lauren B. Resnick, and Kathleen Hammond

Learning Research and Development Center
University of Pittsburgh

1985

Reprinted with permission from G. E. Mackinnon & T. G. Waller (Eds.), Reading research: Advances in theory and practice (Vol. 4, pp. 107-138). Copyright 1985, Academic Press.

The research reported herein was supported by the Learning Research and Development Center, funded in part by the National Institute of Education (NIE), U.S. Department of Education. The opinions expressed do not necessarily reflect the position or policy of NIE, and no official endorsement should be inferred.

LEARNING TO READ: A LONGITUDINAL STUDY OF WORD SKILL DEVELOPMENT IN TWO CURRICULA

ALAN LESGOLD,
LAUREN B. RESNICK, AND
KATHLEEN HAMMOND

Learning Research and Development Center
University of Pittsburgh
Pittsburgh, Pennsylvania

I. Problems in Reading 107
A. Teaching Positions 109
B. Research Issues 110
II. The Need for Longitudinal Data 112
Problems in Longitudinal Studies 112
III. The Present Study 113
A. Children, Schools, and Curricula in the Study 113
B. Measuring Development by Mastery 115
C. Types of Measures Used 115
D. Results 117
E. Structural Modeling Analyses 124
F. Summary and Implications 132
References 137

I. PROBLEMS IN READING

It is often assumed that reading is a very natural skill, one that is readily acquired by anyone who is willing to work hard at learning it. In fact, though, serious efforts to achieve universal literacy are quite novel (Resnick & Resnick, 1977). Many reports of high literacy in some earlier time are based upon low criteria, such as being able to read a fixed Bible passage or sign one's name. Others fail to consider less privileged members of a

ISBN 0-12-572304-0

society in proclaiming the existence of universal literacy. In a sense, reading is a lot like driving a car. If one's expectations of success are low enough, it can appear that the skill is almost universal and that learning it requires nothing but practice. However, adequate driver performance in stressing situations h as driving on ice is not universal. Similarly, reading complex material with understanding is also not universal. Efforts are made to teach effective reading, and those efforts are partly successful. Presumably, if we knew more about how the skills of reading are learned, we could better teach them.

Continuing with our driving metaphor, we might think about how driving is taught. Surely, simple practice of the overall task of driving is a itical aspect of performance. We want drivers to have fully integrated capabilities to keep the car in the correct lane, watch for obstacles ahead, watch for dangers from sides and behind, maintain proper speed, operate windshield wipers, etc. Being able to do each of these things in isolation is not enough, though. Indeed, many of the people who are on the road today know about all of these subskills of driving but fail to coordinate them all into smooth overall performance. On the other hand, there are specific subskills of driving that many individuals simply lack, such as the winter driving skills mentioned above; these could be taught directly with considerable benefit.

One can imagine controversies over whether special subskills of driving should be taught in the context of overall driving practice or whether they should be taught separately. Presumably, most people would favor holistic instruction in driving, and there seems to be no reason why anyone would teach even a critical subskill *completely* outside the context of driving unless this was absolutely necessary. The analogy of this discussion of driving to reading should be apparent. Reading also involves many subskills. Some aspects of reading are exhibited by nearly all members of our society, but overall reading facility is nowhere near universal. Many favor a holistic approach to teaching reading, but, at the same time, there seem to be specific subskills on which performance is often nonoptimal.

The purpose of this article is to review some longitudinal data on the acquisition of reading skill. These data may bear on decisions that are made regarding one particular subskill, rapid word recognition. There is great controversy about whether to teach certain subskills of reading in isolation or as part of general reading practice. One reason for the controversy is that reading builds upon listening. In contrast with driving, where a novice must acquire a complex of mostly novel skills, reading involves the insertion of a new skill, visual word recognition, into a cluster of existing higher-level comprehension skills.[1] It seems quite reasonable that a nondriver would

[1]Of course, there are additional comprehension possibilities which first become feasible when information is being acquired visually rather than aurally.

need to learn *every* aspect of driving, but some believe that a nonreader needs only to learn to recognize words, after which he will, in some sense, be able to read.

In the sections that follow, we review first the alternate teaching approaches used in reading programs today and some elements of psychological theory that bear on these teaching approaches. Then, we discuss the longitudinal research approach we have employed. We believe it can provide useful information relative to several issues of theory and practice. This is followed by a presentation of the method and some outcomes of the study and their interpretation in terms of the issues raised earlier.

A. Teaching Positions

1. Code Approaches

Reading shares many skills with the listening capability that children develop at an earlier age. Everyone learns to listen (barring specific disabilities), but not everyone learns to read. This suggests that much of our effort to teach people to read should focus on the recognition of visually presented words. Words are composed of letters, and those letters occur only in a relatively small number of clusterings, which we might call syllables or spelling patterns. Those of an analytic bent might be moved to break up the task of learning to read according to the following rationale. We can consider reading as consisting of the skills that underlie listening plus word recognition skills. The listening skills are already appearing spontaneously by the time children begin to read, so effort should center on word recognition. The components of words are spelling patterns, so it makes sense to have children learn spelling patterns first and then learn words. Spelling patterns are composed of letters, so children ought to be able to recognize letters before they try to learn spelling patterns.

Such an approach eventually leads to instructional programs that start with subskill training and only later offer substantial experience with the coordinated act of actually reading text. In the context of the initial argument of this article, it might appear that such programs would not be very effective, since they apparently substitute specialized practice for experience with the integrated overall activity of reading. However, a number of meta-analyses of studies comparing reading program effectiveness have found that programs with a phonics component, that is, programs with specific instruction in symbol–sound correspondences, were somewhat more effective than programs which emphasized the integrated activity of reading exclusively (Chall, 1983; Resnick, 1979). Why might this be?

There are several possibilities. First, it may be that certain kinds of word recognition skills are not picked up in the normal course of "just plain reading,'' just as recovery from skids on icy roads is not picked up quickly in the course of "just plain driving." A second possibility is very similar: word recognition practice may be needed in greater amounts than is provided by just reading texts. That is, text reading may not provide very efficient practice for certain aspects of word recognition. Third, the achievement tests used in the studies that contributed to the meta-analyses were, for the most part, primary grades reading tests. These tests are aimed largely at word processing skills, so it is not surprising that they showed an advantage to teaching word processing skills directly.

Finally, it is possible that at the time most of the studies were done, we knew what to tell teachers about how to keep children engaged in reading-relevant word recognition practice, while perhaps we could give only vague guidelines to the teacher using more global teaching approaches. This last possibility deserves elaboration. An instructional system will be effective only if it keeps the learner engaged in an activity that can produce learning. The Chall study and other data have often been taken as indicating that *what* phonics or code-centered reading series teach is the best approximation to what children need to learn. It is also possible that such programs provide no closer an approximation to the optimal than global programs do, but that teachers generally do better at keeping children engaged with phonics materials.

2. *Global Approaches*

Of course, just as code-centered (phonics) approaches appeal to a theoretical viewpoint based upon a task analysis and decomposition of the reading act into component skills, holistic or global methods also appeal to some basic principles, which we have discussed above. Fischer, Burton, and Brown (1978) have perhaps made the strongest case. They point to the problem of learning to ski and note that skiing was formerly taught using a skill decomposition method. Each aspect of skiing was separately learned and practiced. Learning was slow. Then, skis started to be made in a graded series of lengths. Short skis allowed novices to engage in the integrated activity of skiing from the start, without significant risk. Learning became much more rapid. Similar experiences are reported by immersion-type language instruction approaches, such as the Israeli *ulpanim*. A few weeks of continually using a foreign language seem to be more productive than 2 years of classroom instruction that is distributed in small doses and involves subskill practice.

B. Research Issues

Past efforts to compare teaching approaches have too often involved efforts that resembled a horse race. One teacher would try to apply one teaching approach and a second would apply another. Whichever teacher got better results was declared the winner, and his/her aleatoric variation on a vague general principle was taken as evidence that the principle is valid. It is now clear that the choice of curricular approaches cannot be settled using horse-race comparisons. Whichever curriculum "wins" in a particular evaluation may have won for any number of reasons, including management effects, instructional effects, and completeness-of-coverage differences. A more productive approach would be to ask more limited and more controlled questions. In this article, we are concerned with the specific issue of the relationship of word recognition facility to overall comprehension success.[2] If we find, for example, that word recognition facility is an important precursor of achievement in comprehension-related aspects of reading, then there are a number of ways in which curricula might take account of this fact. If, on the other hand, we find that word recognition facility is simply a by-product of overall reading development, that it tracks comprehension achievement but does not precede it, then we will have dealt a strong blow to code-emphasis curricula—although there may still remain management advantages that must be studied separately.

Automaticity Theory

To motivate the specific results we report, we next discuss a particular theoretical viewpoint that is associated with the thinking of some designers of code-emphasis curricula. The longitudinal study we will report in this article was designed partly as a test of theoretical assertions by Perfetti and Lesgold (1977, 1979; Lesgold & Perfetti, 1978) regarding the role of word recognition efficiency in the overall process of comprehending a passage. If, during reading, part of the thinking capacity is given over to word recognition or word understanding, less capacity remains for joining concepts that need to become interrelated in the reader's mind. By automating word recognition and understanding through extensive practice, the mental capacity consumed by those simple processes can be decreased (cf. LaBerge & Samuels, 1974; Shiffrin & Schneider, 1977).

Stated this way, the automaticity argument seems obvious. However, there are two issues worth pointing out. First, there are only limited data to support the importance of word processing automaticity in reading. There are many studies showing that poor readers, at all ages, are slower than better readers at articulating individual words that they see in a visual display (Frederiksen, 1981; Jackson & McClelland, 1979; Perfetti & Lesgold, 1977). However, such correlational evidence must be regarded as relatively

[2]Of course, we also compared two different schools as well as two different curricula. Hence, we caution the reader to watch for qualitative differences in the data of the two groups that can speak to the validity of the conclusions we draw. Simple differences in overall achievement are no more meaningful when we show them than when others do.

weak. Poor readers may be poor because they do not practice very much; more practice might increase word recognition efficiency along with other aspects of reading skill. Second, there are few studies showing that everyday reading stresses the limits of one's processing capabilities, although in extreme cases, such as children who take perhaps 10 to 20 seconds to sound out a word, this seems likely. Put in a slightly different way, common sense and some weak correlational evidence suggest that getting more facile at word processing should be a necessary precursor of getting better at the higher level components of reading, but stronger evidence is needed.

II. THE NEED FOR LONGITUDINAL DATA

A longitudinal study can provide such evidence. By watching students develop their reading skills over time, we can ask the *causal* question: Is facility in word processing an important temporal precursor of improvements in overall reading skill? Of course, the temporal relationship is only one aspect of causal inference, but it is a step closer to where we would like to be. Existing data allow us to make causal inferences about word facility and overall reading achievement of the following very weak form: *the magnitude of A and B are strongly related.* A longitudinal study allows inferences of the form: *changes in A are followed closely in time by corresponding changes in B.* Of course, even in this more advanced form, it is possible that *A* and *B* are both precipitated by an outside cause (*C*), just as spring showers (*C*) produce both wet pavement (*A*) and, shortly thereafter, spring flowers (*B*). But, longitudinal research is a step in the right direction.

Problems in Longitudinal Studies

We set lofty goals for a longitudinal study of the development of reading skill when we started in 1976. We wanted to know how the process of learning unfolds, what some of the critical turning points in the process are, and what aspects of performance during the course of learning to read may be early signals that all is not going well and intervention is needed. We believed that two critical needs faced us in building a usable psychology of reading instruction: (1) to describe successive stages of competence and (2) to account for patterns of transition from one stage to another (cf. Glaser, 1976; Resnick, 1979). Although recent work on the psychology of reading had given us a much richer view of the details of skilled reading performance, we wanted to know more about the *processes of learning to read.*

A first requirement for understanding learning processes is a careful plotting of the actual trajectories of reading skill development in the primary grades. The information then available was based almost entirely on cross-sectional studies, in which children of different ages were compared on some set of reading-related tasks. As discussed above, cross-sectional research designs offer no way of studying individual courses of development, since each child is observed only a single time. In fact, changes in competence are never observed directly in cross-sectional research; instead, changes are inferred from observations made on children of different ages.

Another limitation of most past research on the development of reading skill was that there had been no effort to relate the course of skill development to differences in instruction. Quite typically, all subjects in a reading development experiment had been drawn from a single school or school system, and the instructional program the subjects used was not described in the research reports. We wanted, instead, to deliberately study two different curricula. While this still did not allow controlled observation of the effects of specific instructional components (which could be either differences in what is taught or the capabilities of different sets of students and teachers), it permitted us to discover aspects of learning which varied as a function of curriculum and other aspects which, at least for the range of curricula we examined, did not.

III. THE PRESENT STUDY

We set out to examine how word recognition automaticity develops, and how its development is related to the acquisition of comprehension skill. In particular, our longitudinal design allowed us to go beyond simple correlational analyses to consider the temporal relationships between word automaticity and text understanding skill. We included measures of word recognition speed and oral reading speed in our test battery, along with measures more closely related to comprehension, including the comprehension subtest from a standardized achievement test and tests of semantic judgment about word meanings and the meanings of sentences and paragraphs.

The overall study spanned a period of 5 years, with each cohort of children being followed for 4 years. The core data for each cohort were collected while the children were in first through third grade. Follow-up data were collected for another year in those cohorts where there were enough children remaining in the group to permit sensible interpretation of the data.

A. Children, Schools, and Curricula in the Study

Over this period, we have followed several cohorts of children in two different instructional programs. There was one large cohort in each program, plus smaller cohorts for pilot work and other special purposes. We began the study with over 300 children, to allow for inevitable attrition over the years of children whose families moved away. A total of approximately 80 children remained in the study through third grade.

The *Global-method* cohort children attended school in an urban suburb of 12,000 people, near a major city, with a large proportion of working-class families. Sixty percent were black and 40% were white. The cohort included about 50 children who stayed in the same school for all 3 years of primary reading, and who were still present for fourth-grade achievement testing. Achievement test results confirmed that this group was a representative sample of children for an urban setting. Their mean score on the reading portions of the Metropolitan Achievement Test, Primary Level, at the end of October of first grade corresponded to a grade equivalent of 1.1, a month behind the national average. By October of fourth grade their average scores on the Metropolitan Achievement Test (Total Reading) showed a grade equivalent of 4.3, slightly ahead of the national average.

The school based its primary reading curriculum on the Houghton Mifflin Basal reading program. Although the program was not formally individualized, there was substantial variation in the rate at which children progressed in the curriculum. There were several classrooms at each primary grade level, and there appears to have been some tracking by general ability level in assigning children to these classes. There were also several reading groups within each classroom to accommodate different rates of progress. Thus, most children were receiving instruction at a highly tailored rate.

The *Code-method* cohorts consisted of children in a similar social environment, with closely matching achievement, sex, and race demographics, who were taught using the New Reading System (NRS) developed at LRDC by Beck (Beck & Mitroff, 1972). This curriculum emphasizes word decoding skills along with comprehension skills. It was fully individualized. An important difference that we have noted between the curricula used in our two test sites is that there were explicit criteria in the NRS curriculum for movement from unit to unit. These were based almost completely on which words the children could read, not on their reading speed or any explicit assessment of comprehension. In contrast, children in the Global cohort moved as small groups, based upon teacher decisions about their overall reading performance. There were two Code cohorts. One contained 53 children, of whom about one-third remained after 3 years.[3] The second contained 56

[3]Both schools suffered unanticipated losses of students due to massive unemployment that developed in the area and to uncertainties about court-ordered desegregation plans. We have checked carefully for artifacts that these losses may have produced, and there appears to have been relatively uniform loss of students at all ability levels. Any potential artifacts are noted in the text.

children of whom 13 remained after 3 years. Since the two Code cohorts were in the same school, had the same teachers, and showed similar entering achievement profiles, they were combined for purposes of analysis and reporting. Table I summarizes the schedule for the study.

B. Measuring Development by Mastery

Most developmental studies chart growth as a function of age. In this study, however, the index of development is progress in the reading curriculum, not time. Our thinking was that by examining the performance of students with different levels of ability at points of apparently equal accomplishment, we would be more likely to detect individual differences that are not remediated by simply adjusting the rate of instruction. To this end we established a series of mastery landmarks in each reading curriculum, and each child was tested when he or she reached each landmark. For this purpose the curriculum was divided into segments or levels, each having one or two reading books with accompanying workbooks and related material. Each time a child reached a landmark point in the reading curriculum, he or she was given a battery of tests that assessed word recognition efficiency along with aspects of comprehension efficiency. This strategy allowed us to compare processing capabilities of children who had reached nominally equivalent levels in the curriculum at very different rates.

C. Types of Measures Used

To trace the development of word automation, we used a variety of measures that have proven useful in laboratory studies of the automation phenomenon. These included reaction times for oral reading of individual words and for judgments of word meanings. A further set of measures, used only at the more advanced test points, tested comprehension of sentences and

TABLE I
Schedule of Longitudinal Study

	Cohert and Curriculum		
Year	Code Group B	Code Group C	Global Group D
1977–78	Grade 1	—	—
1978–79	Grade 2	Grade 1	Grade 1
1979–80	Grade 3	Grade 2	Grade 2
1980–81	Follow-up	Grade 3	Grade 3
1981–82	—	Follow-up	Follow-up

passages. To observe how children handled the multiple complexities of meaningful texts as their reading skill developed, samples of oral reading were collected and errors were coded for sensitivity to context and fidelity to the phonemic code. Finally, we included in our database subtest scores from the schools' regular achievement tests, which were administered to all children once a year.

1. Word Vocalization

The most straightforward efficiency measure was word vocalization speed. In this task a word is projected onto a small screen and the child must pronounce the word as quickly as possible. The first appearance of the slide on the screen triggers an electronic clock that runs until the child's voice triggers a voice-operated relay. Responses are recorded on tape. Accuracy scores and mean speed of correct responses were computed for each subject. This measure of word retrieval efficiency was included as an indicator of reading automaticity, on the assumption that as the child becomes able to recognize a word automatically, he or she also retrieves the phonological code more quickly.

2. Semantic Judgments

Children's speed and accuracy in making decisions about the meaning of words were examined in a category matching task. In this task the experimenter said the name of a category (e.g., *animal*), after which a word flashed on the screen. If the word was an instance of the category (e.g., *horse*), the child was to push the *yes* button. If it was not an instance of the category (e.g., *house*), the *no* button was to be pushed. This task was included as a measure of the automaticity of access to word meaning. Many researchers hold that words can be "recognized" in the sense of conveying a concept, even if they are not decoded completely (cf. Perfetti & Lesgold, 1979). That is, one can know what a word means in certain cases before one even knows what word has been seen. While this view is not universal, it seemed appropriate to have a measure of automaticity in processing the meaning of a word.

3. Control Tasks

As a control against the possibility that simple speed of performance might be correlated with reading achievement, simple response time and visual matching tasks were also used.

4. Oral Reading Rate and Error Analyses

The children also read short passages aloud at each test point. We recorded details of each error the child made, as well as the overall time it took to read each passage. The reading speed measures provide an overall index of passage reading efficiency, while the errors can tell us something about the nature of the reading process and especially the interaction of components of that process.

There were two types of passages written for each test level. Familiar passages contained sentences closely adapted from the children's readers and other curricular materials. Transfer passages were less tied to the sentence and phrase structures of the reading materials, and also contained many words (36% of the total) that were used only infrequently (less than 10 times in the reader just finished and less than 3 times in earlier readers). At each test point, the child read aloud one or more familiar passages and one or more transfer passages. Reading errors were analyzed qualitatively, following the procedure of Hood (1975–6) with slight refinements. Reading times were recorded and later converted to speed in words per minute.

5. Standardized Test Data

In the Global-method school, the Metropolitan Achievement Test was administered each year in October. In the Code method school, the Stanford Achievement Test was administered each year in April. We collected these annual standardized test data on all children. In addition, we were, with parents' permission, given any intelligence or readiness test scores the schools had for these children.

D. Results

1. Progress through the Curriculum

A major strategy for data analysis has been to divide each sample of children into three groups (high, medium, and low reading skill) on the basis of second and third grade reading comprehension scores. We can then "look backward" using the longitudinal data to show us the reading development patterns of successful and unsuccessful learners.

The high-skill group contained those children who scored at least one standard deviation above the sample mean on 1 year's test and above the mean on the other, or at least one-half standard deviation above the mean on both. The low-skill group included children who scored at least one standard deviation below the mean on one test and below the mean on the other, or at least one-half standard deviation below the mean on both. About 20–25% of the children were in each extreme group; the remaining children were placed in the medium-skill group.

There was considerable variability in the rate at which children progressed through the curricula. This can be seen in Table II, which shows the earliest and latest grade level at which Levels 3, 5, and 7 of the curricula were completed by the children in each ability subgroup. For the Code

TABLE II
Progress through the Curricula

Group	Range of completion grade levels for					
	Level 3		Level 5		Level 7	
	Earliest	Latest	Earliest	Latest	Earliest	Latest
Global	1.4	2.2	1.9	2.9	2.8	4.0
Low[a]	1.5	2.2	1.9	2.9	2.9	4.0
Medium	1.4	2.2	1.9	2.9	2.8	4.0
High	1.4	1.8	1.9	2.4	2.8	3.0
Code	1.9	3.0	2.1	4.0	2.6	—
Low	2.3	3.0	3.5	4.0	4.0	—
Medium	1.9	3.0	2.4	4.0	2.9	4.0
High	1.9	2.2	2.1	2.9	2.6	3.6

[a]Because different criteria were used, cross cohort comparisons are inappropriate for magnitudes but reasonable for ranges.

group, the fastest high-ability children completed Level Seven 1.4 years ahead of the slowest low-ability children. Some low-ability children did not finish Level 7 even by the end of third grade. In the Global cohort, children advanced through levels with their reading groups, and reading placement was not perfectly correlated with our ability classification. This means that the patterns are less apparent in the table. Nevertheless, the spread between first and last completion of a given level ranged from .8 grade levels at Level 3 (1.4 for the earliest high-ability child to 2.2 for the latest low-ability child) to 1.2 grade levels at Level 7. The low groups progressed much more slowly than did the medium or high groups. However, about 15% of the children in either cohort who were classified as low skill on the basis of their test scores progressed through most of the curriculum at an average or above average rate. One low-skill child moved through the curriculum faster than the average high-skill child! Apparently, some children were able to satisfy their teacher that they were adequately learning the material in their regular lessons; otherwise, they would have been transferred to a slower paced group. Nonetheless, they did poorly on standardized tests later on.

2. *Vocalization Latency*

Results of the word vocalization task are shown in Figs. 1 and 2. Figure 1 shows mean response speeds for correct responses, while Fig. 2 shows mean accuracy levels. Consider the Global group (heavy lines) first. The low and medium Global groups started out taking half a second longer to say a word after it appeared on the screen than did the high group. Some

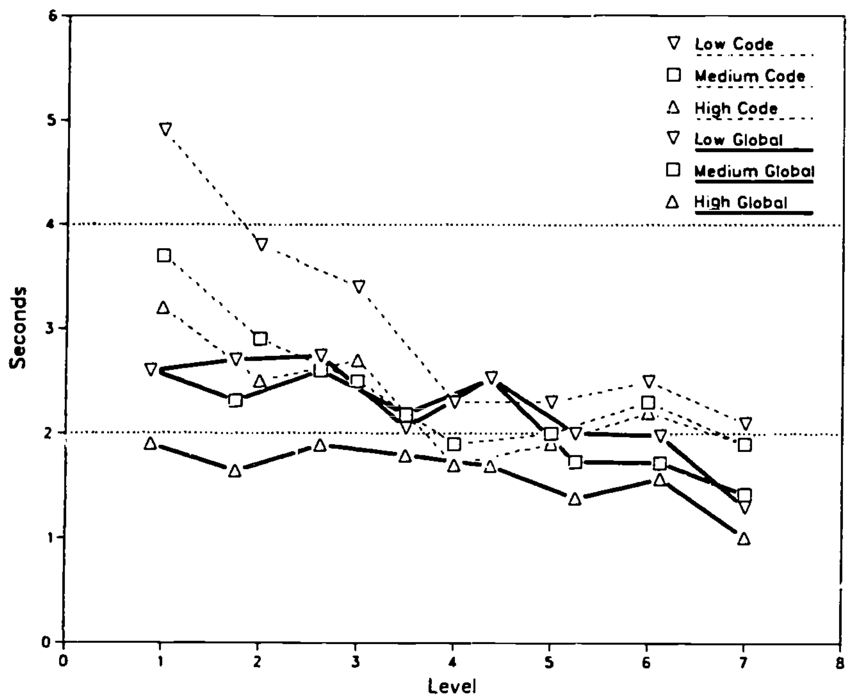


Fig. 1. Vocalization latency results for both cohorts.

difference in speed was maintained over all levels—although all skill groups speeded up by about .8 second by the final test point. There were also accuracy differences in the three groups throughout the study. The low-skill children averaged 40–60% correct, the medium-skill children about 70%, and the high-skill children 80–90%. In the Code cohort, the pattern was similar. An even greater initial speed difference was noted for the Code cohort. The Code children generally were slower at earlier test points than the Global children but showed a more pronounced speeding up as they progressed through the curriculum. In general, the Code students ended up somewhat slower than the Global students and were less accurate.

3. Semantic Judgment Latency

Latencies for the semantic (category) judgment task are presented in Fig. 3. In both cohorts, the low and medium groups took significantly longer than the high group at all test levels, but were not significantly different from each other. Note that all ability levels also showed an improvement in speed over time. Judgment speed dropped from a mean of 4.0 seconds

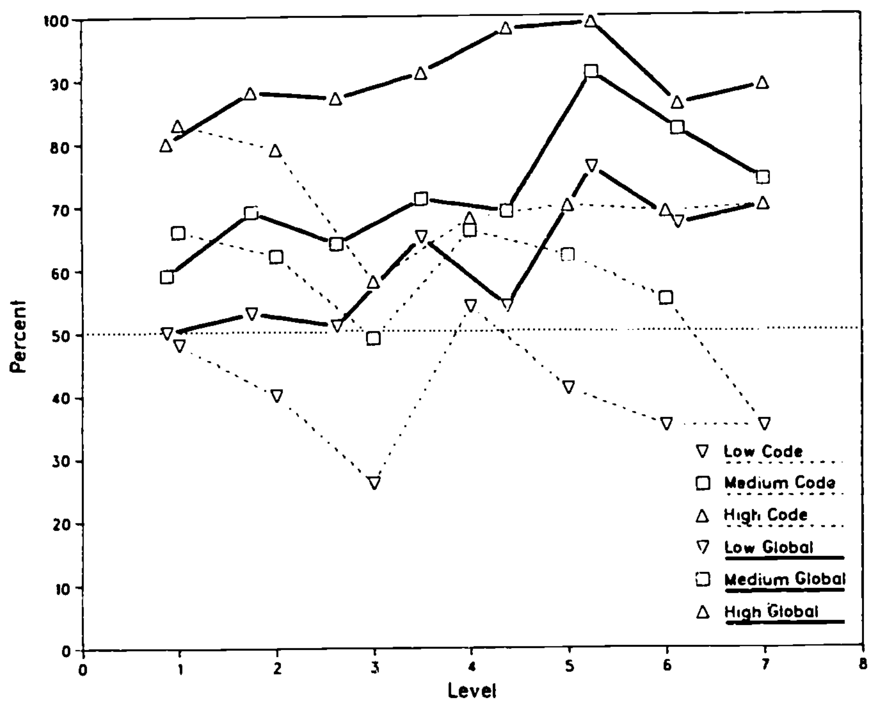


Fig. 2. Vocalization accuracy results for both cohorts.

to a mean of 2.1 seconds from Levels 1 to 7 for the Code cohort, and from 3.0 seconds to 2.0 seconds from Levels 1 to 8 for the Global cohort. The apparent increase in speed for the low-ability Code children at the last test point may be an artifact of sampling: only the fastest low-ability children finished in time to be tested before the end of third grade.

Accuracy of semantic judgments confirms the ability differences. The high ability subjects were more accurate than lower ability subjects. In the Global cohort, there was some rise over time in accuracy. Thus, on the average, these children were becoming both faster and more accurate. In the Code group, by contrast, there appears to have been a speed–accuracy tradeoff. As speed improved over time, accuracy dropped leaving these children at a lower accuracy level than their Global cohort peers at the end of 3 years.

4. Control Tasks

There were no ability-related effects and no significant changes over levels or between cohorts on either the visual matching or the simple reaction time tasks. This means that differences in speed of word vocalization and

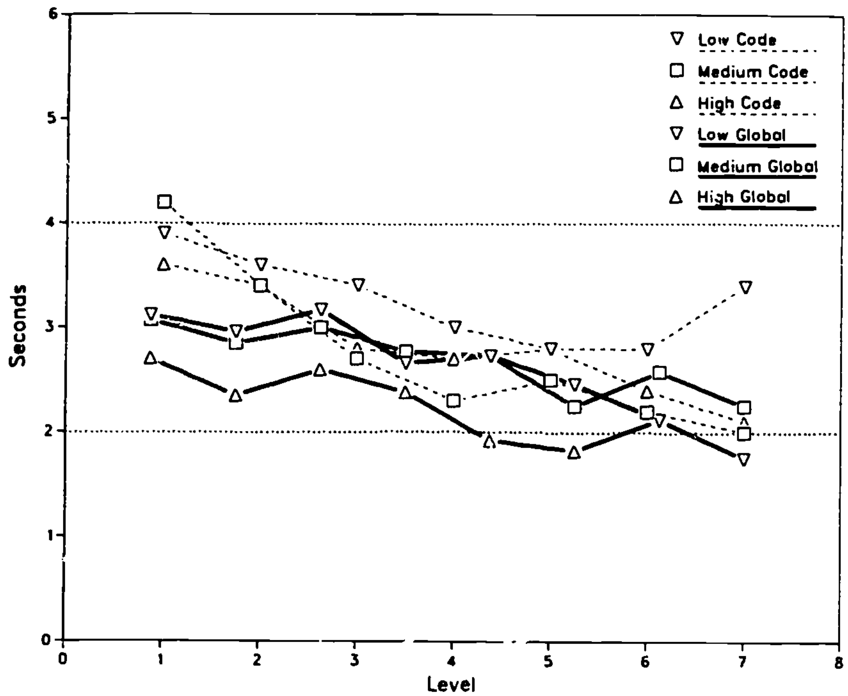


Fig. 3. Categorization judgment latencies for both cohorts.

semantic judgment cannot be attributed to a more general difference in reaction times among subjects.

5. *Oral Reading Speed*

Reading speed data are shown in Figs. 5 and 6. In both cohorts, there were substantial differences between the three ability levels in oral reading rates across the entire period of the study. At the first test point, early in first grade, the high-skill group was reading familiar passages at a rate of about 40 more words per minute than the low group. Although all subgroups gained in speed over levels, a substantial difference was maintained even at the end of third grade. A similar pattern of difference was obtained for the transfer passages, although all groups read them somewhat more slowly. These differences are particularly striking because the children in the lower groups took longer to reach each test point than children in the high-ability groups. They thus had more weeks or even months in which to practice at each reading level. Nevertheless they read more slowly at each level. The low groups remained at less than 30 words per minute on transfer passages through Levels 1 to 4 and increased slowly thereafter. This means that for

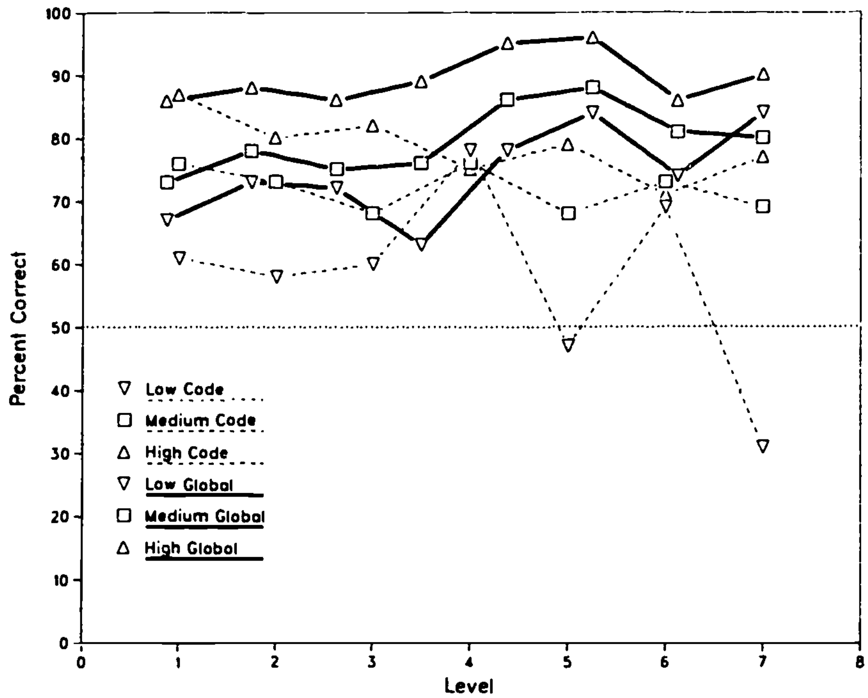


Fig. 4. Categorization judgment accuracy for both cohorts.

about the first 2 years of instruction these children were reading at a rate so slow that comprehension was all but precluded.

Both cohorts ended third grade with mean oral reading speeds between 110 and 120 words per minute on the familiar passages. However, there were substantial differences between cohorts in the early part of the primary curriculum. The Code group began at a mean of 26 words per minute and climbed steadily throughout the three grades. The Global group started at between 65 and 70 words per minute and stayed at that level until late second and early third grade, when they showed a sharp speed-up. It is possible that teachers in the Global classrooms we studied had higher speed/efficiency criteria for progress through the curriculum than did the Code teachers. It is also possible that the overt decoding procedures (such as blending) that *Code* instruction mandates may slow down overall reading speed in the first months of learning to read.

On the transfer passages, shown in Fig. 6, the patterns were slightly different. The Global cohort showed initial performance that was similar to that of the Code group, but ended up slightly lower. Such a pattern might be expected because the initial NRS curriculum provides direct instruction in approaches that will permit decoding of the words in the transfer pas-

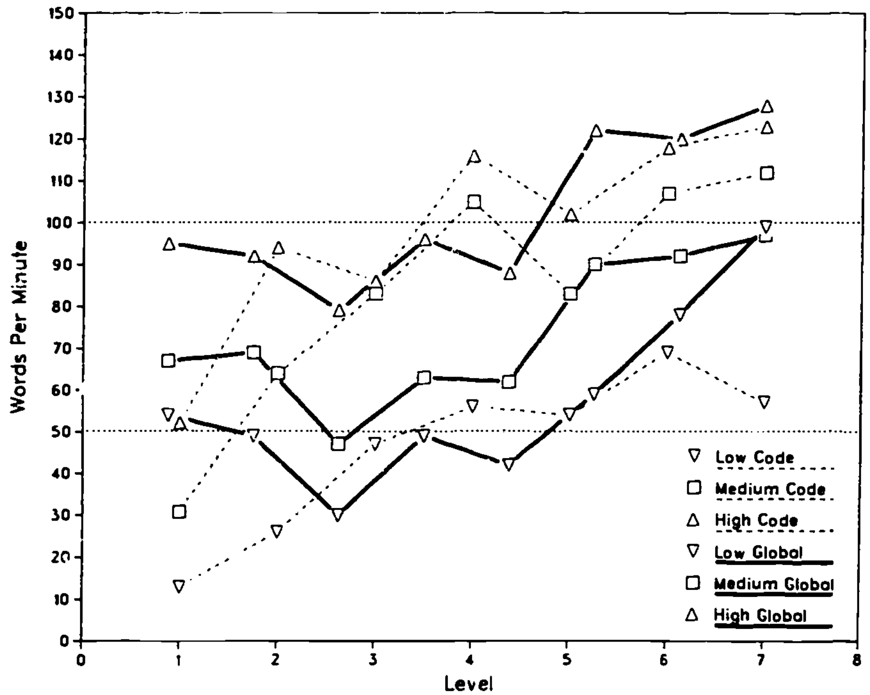


Fig. 5. Oral reading speed for familiar passages.

sages, whereas transfer in nonphonics curricula results from the accumulation (presumably rather slowly) of a broad range of specific word experiences rather than from direct acquisition of decoding rules.

6. *Oral Reading Errors*

There were clear ability group differences in oral reading error rates, as can be seen in Figs. 7 and 8. There were also cohort differences. In the first grade, the error rates on familiar passages in the Code cohort were higher than in the Global cohort; the mean at the first test point was almost 20%, versus 6% for the Global group. However, by second grade both cohorts were showing mean error rates on familiar passages of around 5%. Even the low-ability groups had fewer than 10% errors in the later part of the curriculum. In the Code cohort the low-ability children took longer to improve their error rates on transfer than on familiar passages. Otherwise, the pattern for transfer passages in both cohorts is similar to that for familiar, except that there is a somewhat higher error rate throughout for transfer passages.

Some interesting differences in types of errors were also noted. Qualitative error analysis data are summarized in Tables III and IV. Since no

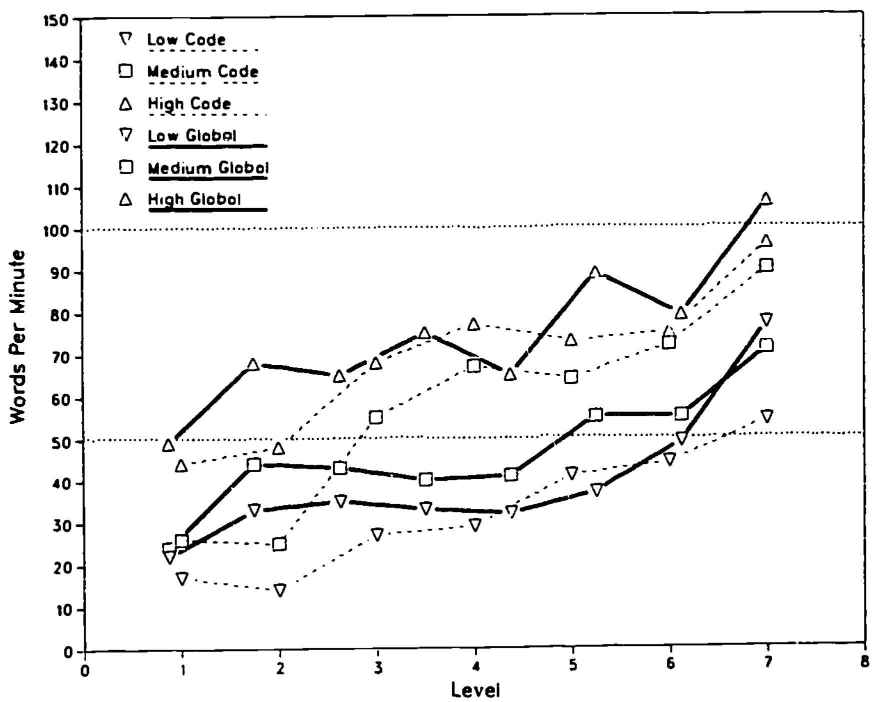


Fig. 6. Oral reading speed for transfer passages.

significant change in patterns across levels was detectable, the oral reading error data were averaged over levels for ease of interpretation. The data show ability by error type interactions. The low-skill children appeared to be intentionally passing over words that they could not recognize or decode (*intentional skip* errors), while the high-skill children seemed to be accidentally missing a few words (*accidental omit* errors). Also, the better readers tended occasionally to insert words as they read, a feature largely absent in the reading of the low-skill children. Finally, high-ability children were more likely than low-ability children to read a word correctly but with the wrong ending. All told, the only striking cohort difference is that Code Cohort children had somewhat more nonsense errors than the Global children.

E. Structural Modeling Analyses

The general pattern of the results discussed above suggests that individual word processing skill is an important component of high levels of reading achievement in the primary grades. This is in accord with previous findings

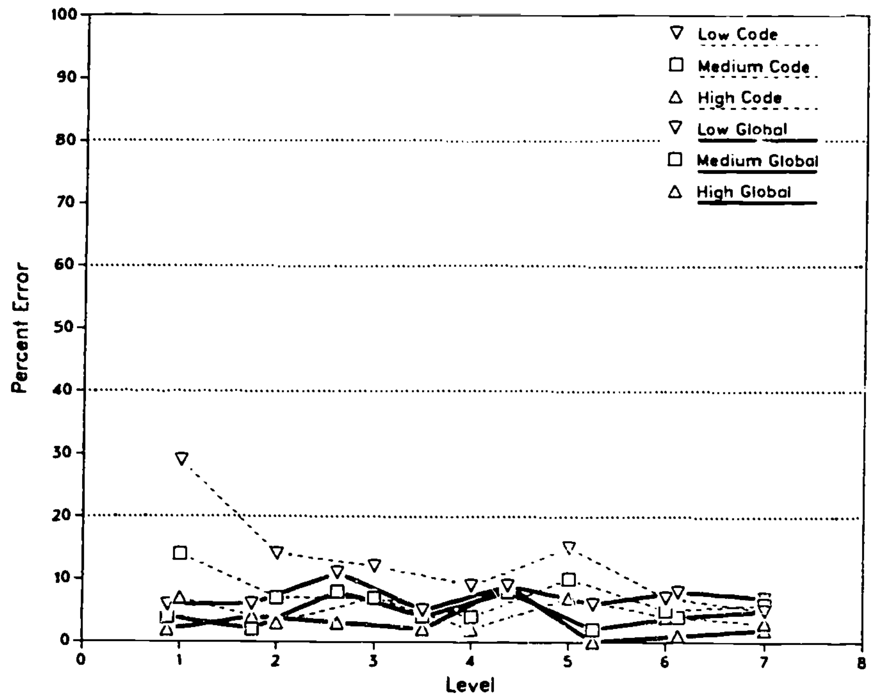


Fig. 7. Oral reading error rates for familiar passages.

TABLE III

Qualitative Analysis of Oral Reading Errors, by Skill Levels: *Code* Cohort

Error type	High skill	Medium skill	Low skill
Child stop reading for 5 seconds or more	7.6	6.8	5.6
Child *accidentally omits* a word	6.4	3.9	4.0
Child *intentionally skips* a word	2.6	5.3	9.9
Extra word inserted	4.1	2.9	1.9
Word order switched	.2	.4	.2
Letter reversals within word (*was* for *saw*)	.7	.5	1.0
Correct word with wrong ending	12.3	10.4	9.3
Wrong word with correct ending	5.1	5.4	4.3
Other word substitutions	49.9	54.9	51.1
Nonsense (nonword)	11.0	9.5	12.7

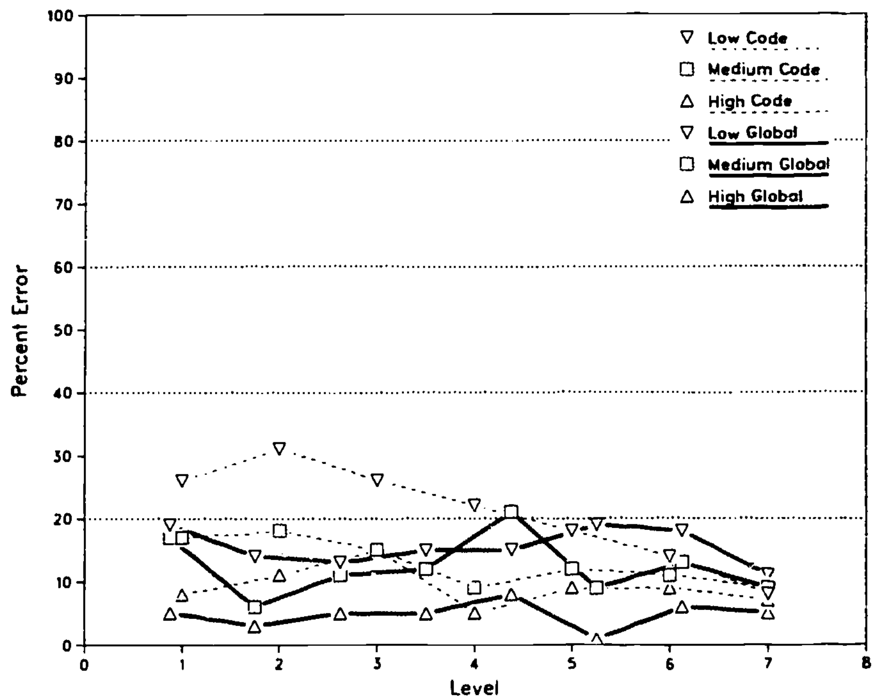


Fig. 8. Oral reading error rates for transfer passages.

TABLE IV
Qualitative Analysis of Oral Reading Errors, by Skill Levels: *Global* Cohort

Error type	High skill	Medium skill	Low skill
Child stop reading for 5 seconds or more	5.8	6.2	5.7
Child *accidentally omits* a word	13.0	5.4	5.8
Child *intentionally skips* a word	2.2	9.0	15.0
Extra word inserted	5.0	2.5	1.5
Word order switched	.1	.3	.0
Letter reversals within word (*was* for *saw*)	1.5	.6	1.0
Correct word with wrong ending	12.8	12.2	8.0
Wrong word with correct ending	6.4	5.8	5.6
Other word substitutions	46.5	49.6	51.2
Nonsense (nonword)	6.7	8.4	6.3

on the relations between word processing automaticity and reading skill reported in the literature. Our interest, however, was in closing the gap between correlational data and the causal hypotheses we would like to confirm. We knew from prior research, as well as from the data in the present study, that children with reading comprehension difficulties lack individual word skills, especially automaticity. But given only the correlations of any two measures taken at about the same time, it is unwise to draw causal conclusions. For example, we cannot decide from such data whether the word recognition facility is necessary for comprehension to succeed or whether it is the result of exercising a higher level of comprehension skill (i.e., perhaps good readers read more and thus become faster word processors).

Longitudinal data with multiple measurements on the same subjects at separate times can permit stronger tests of the hypothesis that word processing facility is an important precursor of reading facility. In examining data this way, one looks for associations between a variable measured early in the time sequence and one measured later. Since early occurring events can cause later ones, but later events cannot cause earlier ones, the causal relations between word skill and comprehension can be inferred by comparing the extent to which early word skill automaticity predicts later comprehension with the extent to which comprehension predicts later automaticity. If one of the associations is reliably greater than the other, a primary direction of causality can be inferred.[4] This is the basic logic underlying a number of longitudinal analysis methods, including cross-lag panel analysis (Campbell, 1963; Kenny, 1975) and structural equation modeling (Joreskog & Sorbom, 1978).

We originally analyzed our data using structural equations modeling procedures (Joreskog & Sorbom, 1978). For example, Lesgold and Resnick (1982) specified a model in which constructs representing word processing speed at different points in the curriculum were estimated from measures of word recognition speed and oral reading speed. For the Global cohort (the only one on which enough data were available at the time), the best fitting models of the correlations among these word processing speed measures and reading comprehension scores showed larger predictive paths from *early word processing to subsequent comprehension* than vice versa. For example, the average weighting for paths from speed to subsequent comprehension was as great as the average path from 1 year's comprehension

[4]If relationships in both directions are significant (or not significantly different), however, causal inferences cannot easily be made. Further, while longitudinal data can test specific causal models, only an intervention study in which the presumed causal skill is trained and the predicted effect produced can firmly validate the causal prediction.

to the next year's, while the average path from comprehension to subsequent speed measures was only one-tenth as large. Lesgold and Resnick concluded that during the first 2 years of schooling, word processing speed is an essential precursor of comprehension success.

To the extent that these results hold up over the full 3-year time period, they may be taken as implying a causal relationship between efficiency of word processing and overall comprehension competence. The present report provides a first look at the data from the entire 3-year period, for both cohorts. Basically, the data support the viewpoints taken in our earlier paper, in that word processing measures predict later reading comprehension performance better than vice versa. After considerable reflection, however, we have switched to a different method of path analysis than that used earlier. Before discussing the specifics of our findings, we will explain the reasons for this new approach.

1. Uniqueness of Predicted Variance

The structural equations approach is one in which a best overall fit of path parameters is estimated. Reliable and accurate path analyses can emerge from this modeling method only when the structure variables are relatively nonoverlapping in their sources. If two structure variables overlap in the capabilities they measure, then there will be some instability in the path weightings generated. Whichever of the two has the most unique predictive power will likely have heavy path weights to the criterion variable, even if the common predictive ability of the two is very high. Because our measures of word processing accuracy and word processing facility were known to overlap highly, we feared that the predictive capability they shared would automatically accrue to whichever of them had a slight edge in unique predictive power. In fact, this happened when we performed structural equations analyses of the final data. There was the appearance of strong differences between curricula in the relative importance of word processing *accuracy* and word processing *speed* to ultimate reading achievement.

To avoid misinterpretation of this result, we have searched for approaches that would allow us to make very clear the extent to which apparent path differences were due to shared common variance being allocated to whichever predictor had the edge in unique predictability. The approach we developed is one that is also suggested by some statisticians (e.g., Marascuilo & Levin, 1983). Specifically, we establish path weights using multiple regression techniques, and then we perform commonality analyses (Kerlinger & Pedhazur, 1973) whenever the extent to which variables share predictive power is an issue. With this approach in mind, let us turn to the analyses of our data.

2. *Path Analysis Results*

The basic path analysis results are shown in Figs. 9 and 10. In these analyses, we created several derived variables to use as predictors. These measures were created by converting raw measures to z-scores and then averaging over multiple measures and multiple testing periods. The *Speed* measures represent an averaging of four measures: word vocalization response speed, category judgment response speed, oral reading speed for familiar passages, and oral reading speed for transfer passages. The *Accuracy* measures represent an averaging of word vocalization and category judgment accuracies. Data were also collapsed over adjacent test points.

The *Early* measures come from Levels 1 and 2 of the Code Cohort and Levels 1 through 3 of the Global Cohort, both spanning most of the first grade. Similarly, the *Middle* measures were averaged over two levels subsuming approximately the second grade, and the *Late* measures were averaged over the last three Levels, subsuming third grade for an average student. The reading comprehension subtests of the standardized achievement tests given by the schools were also used in these analyses. Appropriate achievement test data were available only beginning in second grade for the Code cohorts, and the number of students surviving until the fourth grade test for that group was too low to include fourth-grade data (about 15 children). Finally, the vertical dimension of the figures approximately represents the passage of time over the 3 years of primary instruction.

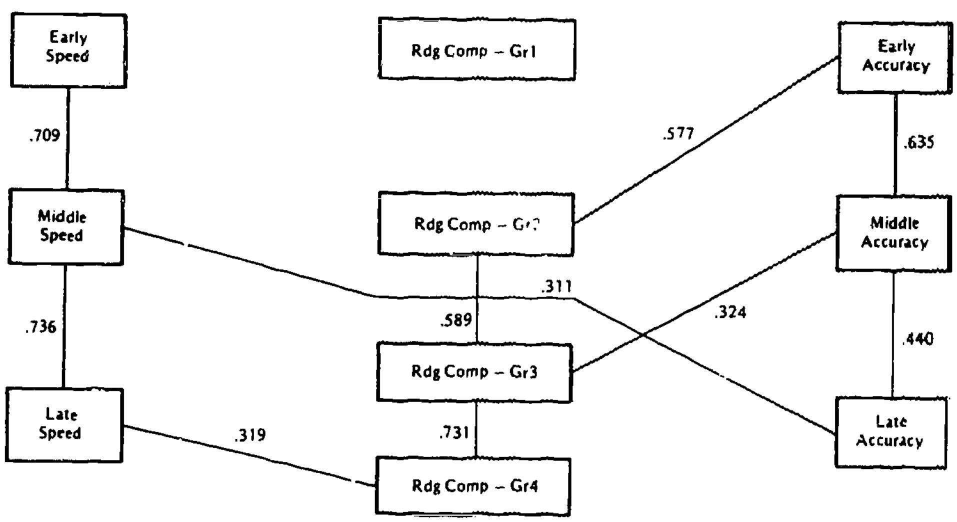


Fig. 9. Path analysis for *Code* cohort.

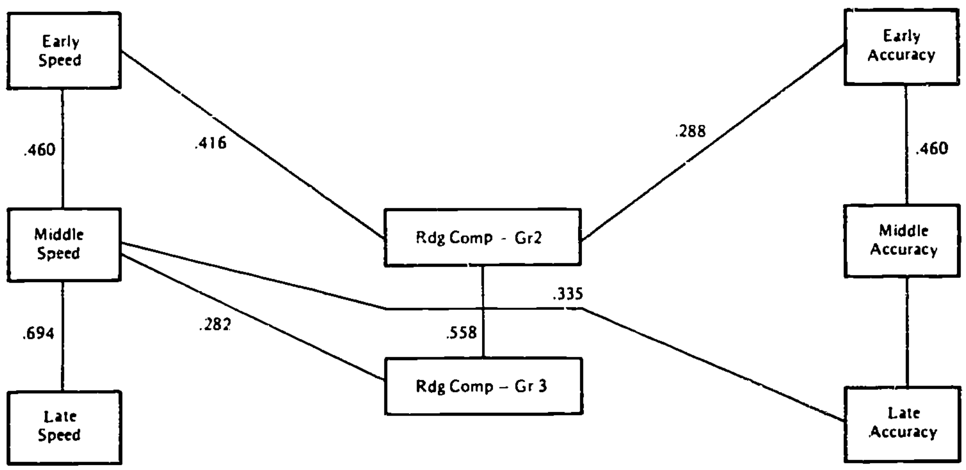


Fig. 10. Path analysis for *Global* cohort.

The results confirm our expectations that *Speed* and *Accuracy* measures would be better predictors of later reading comprehension performance than vice versa. Also, there is an interesting effect in both cohorts, with *Middle Speed* predicting *Late Accuracy* rather strongly. This may be due to the fact that the shorter words learned earlier in the curriculum contain spelling units that must be handled with facility in order to accurately process longer words that appear later. Finally, there is the appearance that progress in the Code cohorts is driven primarily by progress in word processing speed, while progress in the Global cohort seems more to be related to word processing accuracy. This apparent difference is what drove us to use commonality analyses on our data, so we turn to those analyses next.

Figures 11 and 12 show the results of commonality analyses of the regression of second-, third-, and fourth-grade reading comprehension on the Speed, Accuracy, and standardized reading comprehension test measures from the year before the predicted performance.[5] The numbers in the Venn diagram circles represent proportions of the variance of reading comprehension accounted for by each of the predictor variables uniquely and in combination with others.

A first result to note is that for the Global cohort most of the predictive power is shared. This is less marked for the Code cohort. Another differ-

[5]First-grade reading comprehension tests were not given at the same time to all Code-Group students, so we could not sensibly use first-grade reading comprehension test scores as a predictor for that cohort. We also did not retain enough fourth-grade code children to enable any regression analyses for fourth-grade Code group reading comprehension ability.

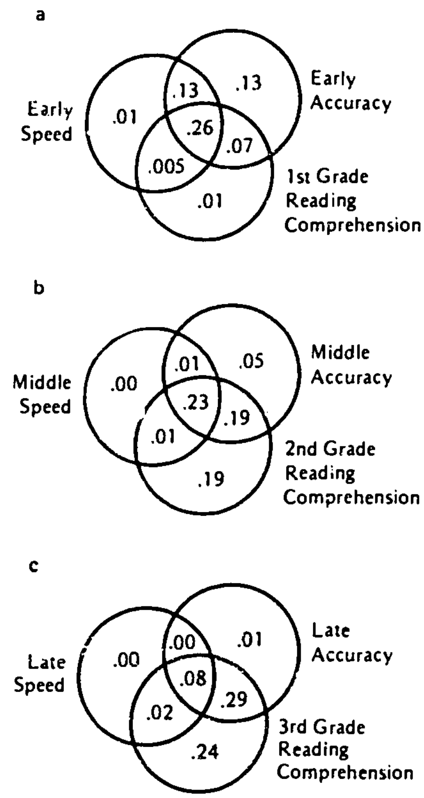
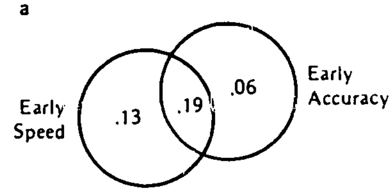


Fig. 11. Commonality analysis of Global group predictors of reading comprehension in (a) second grade, (b) third grade, and (c) fourth grade.

ence between cohorts is in the small amount of unique predictability that comes from *Speed* versus *Accuracy*. The Global cohort shows some predictive power for accuracy and none for speed, while the pattern is reversed for the Code cohort. A third important result is that the contribution of 1 year's reading comprehension to predicting the next year's increases over time, while the contribution of word-processing measures declines. This replicates an earlier study by Curtis (1980). We will say more about the decreasing importance of individual word processing measures in predicting achievement in later grades later in this article.

We also performed commonality analyses using effective reading speed as the criterion variable. Effective reading speed was computed by averaging the *z*-scores for oral reading speed and oral reading accuracy. The predictor measures were word vocalization efficiency (the average of the a



b

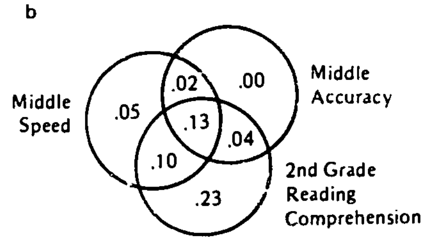


Fig. 12. Commonality analysis of Code group predictors of reading comprehension in (a) second grade and (b) third grade.

vocalization speed and accuracy z-scores) and semantic processing efficiency (the average of the category judgment latency and accuracy z-scores). Figure 13 shows the results for the Global group, while Fig. 14 shows the same results for the Code group. At both second and third grade, the prediction of effective reading speed derives from variance shared by all predictors in the case of the Global cohort while there is unique predictive variance in the Code Cohort. Another difference between the groups is in the contribution of semantic access efficiency versus word vocalization efficiency. Vocalization efficiency carries more weight in the Global cohort predictions while semantic access efficiency carries more for the Code group.

F. Summary and Implications

The longitudinal data we have gathered show a clear relationship between word recognition efficiency early in learning and reading comprehension performance later on. This is consistent with the view that efficient word recognition skill that does not require substantial allocation of limited cognitive resources is important to the overall development of reading skill. The asymmetry of the relationship—the fact that early comprehension skill is not associated with later word recognition efficiency—strongly suggests that word-level skills facilitate the acquisition of comprehension skills. While certain alternative possibilities, such as another unidentified factor that first causes word recognition and later comprehension skill, are not strictly ruled out, our findings do require that we reject the idea that word recognition

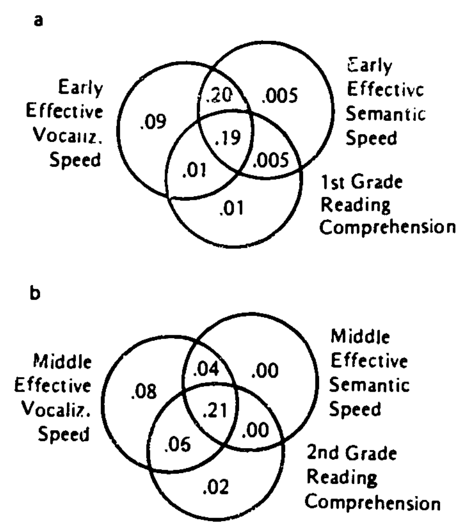


Fig. 13. Commonality analyses for predictions of Global Group effective reading speed in (a) second grade and (b) third grade.

facility develops as a direct *result* of comprehension skill. It seems, then, that developing word recognition efficiency ought to be an important goal of early reading instruction.

Our data suggest that the goal of word processing efficiency was not being

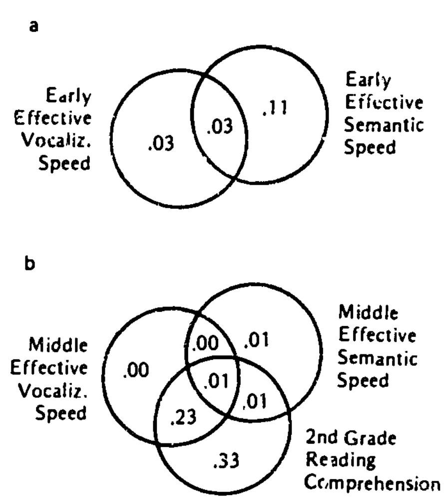


Fig. 14. Commonality analyses for predictions of Code Group effective reading speed in (a) second grade and (b) third grade.

fully achieved in either of the classrooms we studied. It is clear that a certain proportion of students (perhaps 15%) in each cohort was allowed to pass from unit to unit of the reading program without developing word processing facility. This was perhaps to be expected in the Global method classrooms, which did not have formal criteria for passing from one unit to the next. However, it is, on first consideration, surprising for the individualized Code method classrooms. There, to pass from one level to the next each child had to pass a test demonstrating ability to read words that used the spelling patterns that had been taught. How could it be that so many children who could read words accurately enough to pass these successive tests had, at the end of third grade, word recognition skills as weak as those we have documented for the low-ability group? The answer seems to lie in exactly what was tested, and therefore emphasized, in instruction. In the code program studied children were taught a phoneme-by-phoneme decoding and blending algorithm and were encouraged to use it whenever they did not immediately recognize a word. In the tests, and in teaching, children were given as long as they needed to pronounce individual words, using the algorithm as necessary. The emphasis, in other words, was on accurate but not on automatic word recognition. Under these conditions, some children were able to pass from level to level despite very slow word recognition.[6]

The Code method's strong focus on the accuracy component of word recognition may also account for the different patterns that the two cohorts displayed in the structural equation and commonality analyses. In both cohorts, word facility developed ahead of and was related (perhaps causally) to later comprehension skill. However, in the Code cohort *speed* of word recognition carried the (slightly) greater weight in these relationships, whereas in the Global cohort *accuracy* of recognition carried the greater weight. A consideration of the teaching methods and what they imply for acquisition suggests why this difference should occur. In the Code method classrooms, some of the individual components of reading skill were separated out and given extra attention and practice. Of the many components of word recognition skill, the one that instruction emphasized was recognition accuracy. Under these conditions, accuracy could be expected to develop ahead of other components, and perhaps not be fully integrated into the child's general reading competence, at least for weaker readers. Under these conditions, accuracy of word recognition cannot be expected to index how well early reading skill was developing and would not forecast later

[6]In fact, the accuracy criterion was not all that high, either. Had a higher accuracy criterion been used, the practice needed to produce high accuracy would have forced automation of word recognition, but at the expense of less time spent on the higher level components of reading.

comprehension ability. Speed of word recognition, which was not explicitly stressed in instruction, can and does do this job.

In the Global method classrooms, by contrast, instruction was organized around the reading of successive stories. In this context, children discussed and briefly practiced various specific elements of reading: word meanings, sentence meanings, grapheme–phoneme correspondences, related "world knowledge," and the like. None of these components took precedence over any other; instead all were expected to develop simultaneously, more or less as a "package." Under these conditions, accuracy can be expected to serve as a good index of how reading skill is developing. One reason for this is that accuracy of word recognition is very easy for teachers to assess, even without formal tests. It is therefore likely to play an important role in which reading groups children are assigned to. These assignments in turn determine how much and what kind of reading exposure children are likely to have. A higher reading group means reading more different stories and thus more practice on all components of reading. Accuracy was also more reliably measured than speed in our study, since there is substantial variability in reaction time data, especially for young children. This, too, would favor the emergence of accuracy as an indicator of how general reading skill was developing.

The commonality analysis results for the two cohorts support the interpretation that development was occurring globally in one cohort and componentially in the other. For example, in the predictions of second grade reading comprehension scores for the Global cohort, the biggest source of predictive variance was common to earlier word recognition speed, accuracy, and overall reading comprehension. For the Code cohort, more of the predictive variance was unique to the individual variables or common only to pairs of the three variables. There was a similar pattern for predicting third grade reading speed. This is even more marked in predicting effective reading speed (Figs. 13 and 14). Predicting the Global cohort's third-grade effective reading speed, for example, the bulk of the explained variance was again shared (.21), while there was essentially no triply shared variance for the Code cohort (.01). A final supporting fact is that in the Global cohort, speed and accuracy of semantic category judgments improved together over time (Figs. 3 and 4), while in the Code cohort accuracy dropped as speed increased.

We also found that the two different teaching approaches resulted in substantial qualitative differences in trajectories for acquisition of some aspects of word recognition facility. The most striking difference is that in all speed measures, the Code children began more slowly, and then showed sharp increases in speed; while the Global children began faster and showed a very gradual increase over the 3 years of the study (see Figs. 1, 3, and 5).

This difference in shape of the trajectory is not surprising: the Code children were deliberately slowing down during the early units of instruction in order to use the blending algorithm. By the end of third grade, the oral reading differences had disappeared. Indeed, there is some indication that middle-ability children in the Code group may even have surpassed their Global-group peers in third grade, having started out behind them. On the isolated-word tasks, the Global group showed a small persisting advantage, ending a half second faster in word vocalization and slightly faster in the category judgment task.

It is not yet entirely clear how these cohort differences ought to be interpreted. We have collected some further data on those children in our samples who remained in the same schools into fourth and fifth grades. These data will allow us to relate the early patterns described here to overall reading competence, particularly in comprehension, later in elementary school. Analysis of these data will perhaps give a clearer sense of the meaning of the different developmental patterns that we have observed in the present study. For the moment, we can only say that there is evidence that there are indeed multiple pathways to the goal of reading fluency and that instructional programs can influence which of these pathways a child takes.

While our present findings do not give us a basis for choosing between Global and Code approaches to teaching beginning reading, they do suggest that neither approach—at least as practiced in the classrooms we studied—was providing the strongest possible support for developing word recognition efficiency. It is clear that children in the low-ability groups of both cohorts were about to leave the primary grades with reading skills inadequate to the demands likely to be placed on them in subsequent schooling. Reading speeds of 50 to 70 words per minute—the average speed on transfer passages for the low-ability groups—are so slow as to interfere with comprehension even of easy material, and are certainly unlikely to leave much memory capacity free for *developing* new comprehension abilities.

Some of this problem could probably be resolved by better management of reading instruction, in particular resisting the temptation to pass children on to higher levels in the curriculum when fundamental abilities were still poorly established. However, to hold children back in order to do more of the same kind of teaching is unlikely to produce dramatic successes. Instead, ways need to be found to build word recognition efficiency more effectively. Certainly, radical alterations of curriculum should not be proposed in the absence of direct tests of the proposed changes. However, the data in hand now suggest that combing the curriculum for opportunities to increase efficient word recognition and the ability to deal quickly with word meaning will be a worthwhile enterprise, especially if it is combined with extensive practice in reading real texts for meaning.

ACKNOWLEDGMENTS

The research we report was funded by the Learning Research and Development Center, University of Pittsburgh, through an institutional grant from the National Institute of Education. The ideas and positions in the article are wholly the authors', and no endorsement by NIE or any other government agency of the contents of this article is implied. Deborah Wijnberg, Hope Cordonier, and Carol Sharp were the primary research assistants for this project, and Mary E. Curtis participated substantially in the design of the study and the first years of implementation.

REFERENCES

Beck, I. L., & Mitroff, D. C. *The rationale and design of a primary grades reading system for an individualized classroom.* Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1972/4.

Campbell, D. T. From description to experimentation: Interpreting trends as quasi experiments. In C. W. Harris (Ed.), *Problems in measuring change.* Madison: Univ. of Wisconsin Press, 1963.

Chall, J. S. *Learning to read: The great debate.* New York: McGraw-Hill, 1983.

Curtis, M. E. Development of components of reading skill. *Journal of Educational Psychology,* 1980, 72, 656–699.

Fischer, G., Burton, R. R., & Brown, J. S. Aspects of a theory of simplification, debugging, and coaching. *Proceedings of the second national conference of the Canadian Society for Computational Studies of Intelligence.* Toronto: Univ. of Toronto Press, 1978.

Frederiksen, J. R. Sources of process interactions in reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, New Jersey: Erlbaum, 1981.

Glaser, R. Components of a psychology of instruction. *Review of Educational Research,* 1976, **46,** 1-23.

Hood, J. Qualitative analysis of oral reading errors: The inter-judge reliability of scores. *Reading Research Quarterly,* 1975–76, **11,** 577–598.

Jackson, M. D., & McClelland, J. J. Processing determinants of reading speed. *Journal of Experimental Psychology: General,* 1979, **108**(2), 151-181.

Joreskog, K. G., & Sorbom, D. *LISREL: Analysis of linear structural relationships by the method of maximum likelihood* (User's guide, Version IV, Release 2). Chicago: International Education Services, 1978.

Kenny, D. A. Cross-lag panel correlation: A test for spuriousness. *Psychological Bulletin,* 1975, **82,** 887–903.

Kerlinger, F. N., & Pedhazur, E. J. *Multiple regression in behavioral research.* New York: Holt, 1973.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology,* 1974, **6,** 293–323.

Lesgold, A. M., & Perfetti, C. A. Interactive processes in reading comprehension. *Discourse Processes,* 1978, **1,** 323–336.

Lesgold, A. M., & Resnick, L. B. How reading disabilities develop: Perspectives from a longitudinal study. In J. P. Das, R. Mulcahy, & A. E. Wall (Eds.), *Theory and research in learning disability,* New York: Plenum, 1982.

Marascuilo, L. A., & Levin, J. R. *Multivariate statistics in the social sciences: A researcher's guide.* Monterey, California: Brooks/Cole, 1983.

Perfetti, C. A., & Lesgold, A. M. Discourse processing and sources of individual differences. In P. Carpenter & M. Just (Eds.), *Cognitive processes in comprehension*. Hillsdale, New Jersey: Erlbaum, 1977.

Perfetti, C. A., & Lesgold, A. M. Coding and comprehension in skilled reading. In L. B. Resnick & P. Weaver (Eds.), *Theory and practice of early reading*. Hillsdale, New Jersey: Erlbaum, 1979.

Resnick, L. B. Theory and practice in beginning reading instruction. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 3). Hillsdale, New Jersey: Erlbaum, 1979.

Resnick, D. P., & Resnick, L. B. The nature of literacy: An historical exploration. *Harvard Educational Review*, 1977, **47**(3), 370–385.

Shiffrin, R. M., & Schneider, W. Controlled and automatic human information processing (Vol. 2). Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 1977, **84**, 127–190.