A LONGITUDINAL STUDY OF THE EFFECTS OF THE *BRIDGE* READING PROGRAM FOR CHILDREN AT RISK FOR READING FAILURE

Andrew Biemiller and Linda S. Siegel

Abstract. **The effectiveness of an experimental reading program, *Bridge*, was compared to "Whole Language" reading instruction in two schools in economically disadvantaged areas. The *Bridge* program uses icons or picture symbols to facilitate print word identification. Children who were predicted to be at risk for reading difficulties comprised the *Bridge* (N = 42) and Whole Language (N = 64) samples. Predictions of at-risk reading status were made using a test of phonemic analysis skills (the *Lindamood Auditory Conceptualization Test*), Denkla and Rudel's *Rapid Letter Naming Test*, and a test of syntactic awareness. The *Bridge* group performed significantly better on the WRAT Reading test than the Whole Language group at the end of grade 1 and grade 2. The *Bridge* program appears to be useful for children who are at risk for reading failure.**

There has been a growing agreement among reading researchers that a crucial step in learning to read consists of mastering the identification of printed words (Gough & Tunmer, 1986; Juel, 1988; Stanovich, 1986; Vellutino, 1991). Juel, Griffith, and Gough (1986) and Byrne, Freebody, and Gates (1992) provide clear evidence of the importance of word identification for the development of reading comprehension.

At the level of identifying single words, reading failure generally involves difficulties in (a) identifying printed words accurately and (b) reading words rapidly enough to allow comprehension of sentences. These two problems are usually correlated, and are related to cognitive development (Biemiller, 1977-78; Blachman, 1984; Cantalini, 1987; Lovett, 1987; Perfetti, 1985; Stanovich, 1986; Stanovich, Nathan, & Zolman, 1988). We refer to these as "reading-specific skills."

Language difficulties are also related to reading success (Perfetti, 1985; Siegel & Ryan, 1988; Stanovich, 1986; Vellutino, 1979). The experimental curriculum used in the present study was designed to provide opportunity to ensure understanding of all vocabulary used.

In addition to reading-related skills and language difficulties, several theorists (Gleitman & Rozin, 1972; Mounoud, 1986; Siegel, in press) have suggested that when "reading" involves converting printed words to speech and then extracting meaning from speech, the load on some young children's working memories may simply exceed their capacities at that moment in their development. In this project, we used an approach to beginning reading that was intended to reduce the cognitive load in working memory on word identification.

Developmentally appropriate initial reading instruction for children who have not yet reached the necessary level of working memory must somehow bypass the need for relying exclusively on decoding print to sound and then sound to

ANDREW BIEMILLER, Ph.D., is professor, Ontario Institute for Studies in Education, The University of Toronto.

LINDA S. SIEGEL, Ph.D., is professor, Department of Educational Psychology and Special Education, University of British Columbia.

meaning; and must control the complexity of written syntax and vocabulary with which children must cope. Such a program should permit children who are less cognitively advanced in terms of working memory to begin reading instruction without initial failure. Similarly, ESL (English as a Second Language) children may benefit from controlled vocabulary loads and reduction of the cognitive demands of word identification processes.

Dewsbury, Jennings, and Boyle (1983) have developed a reading program that meets these conditions. Their *Bridge* program begins reading instruction with the use of simplified visual pictures or "icons" of nouns (initially various foods, vehicles, and animals), acts (e.g. "find," "like," "circle," etc.), and adjectives (e.g., "big," "little"). These visual signs are then assembled into simple sentences which the children both read and write. Printed words are initially paired with icons. Exercises from the first lesson on gradually ease children into using words represented only in print. The syntactic complexity of the language the children read is carefully controlled. (Examples are given in Figure 1.) The objective of using icons and normal, but simple language ensures success at deriving meaning from written language. Very simplified content is used, along with workbooks and games (e.g., Lotto, Concentration) designed to facilitate the transition from icons to print.

Text from book 1:

my banana

Text from book 29:

they have a plane

Text from book 50 (no icons)

She jumps on

Baby Bear's bed.

"I like this bed," she says.

"It is good."

Figure 1. Sample sentences from books at different points in the *Bridge* program.

The underlying assumption of the program is that the icons or picture symbols can be linked to the child's lexicon much more readily than print symbols. These icons then serve to help the child learn the print symbols over a number of successful trials in reading and activity contexts. The main goal of the program is to achieve a sight vocabulary for print of approximately 150 words. This target vocabulary is based on the Dolch list of familiar words (Dolch, 1936).

We adapted the published *Bridge* program to render it more usable in regular classrooms. Our adaptation included preparation of some materials described in the program but not supplied, preparation of short lesson plans for every lesson, and development of some additional workbooks and games. The intended effect of this introduction to beginning reading was that when the child had matured sufficiently to profit from normal reading instruction, she or he would have a reading vocabulary of 150 sight words and a considerable sense of success at reading activities.

DESIGN AND METHODS

Design

The research involved an experimental design with five Bridge classrooms and seven "Whole Language" classrooms. Initially, grade 1 teachers in two schools serving highly disadvantaged children were invited to try the Bridge program in their classes with children whom they believed would profit from the program. In year 1, three teachers with grade 1 children in one of the schools agreed to do so. These teachers had three, seven, and 15 years of teaching experience. The Bridge treatment was replicated in year 2 in two first-grade classrooms in the same school. Two new first-year teachers were involved in this replication.[1] Thus, a total of five classrooms were in the Bridge treatment.

The "Whole Language" population consisted of children in the same two schools as the seven teachers who had grade 1 children in year 1. All of these teachers reported that they preferred to use the "Whole Language" methods encouraged by the board of education. These teachers ranged from one to 20 years of teaching experience, averaging seven years.

Implementation of the board-recommended "Whole Language" program varied considerably from classroom to classroom, as there was no general curriculum guideline. (For discussion of the problem of describing Whole Language instruction, see Edelsky, 1990; McKenna, Robinson, & Miller, 1990; and Stahl & Miller, 1989.) However, the programs avoided use of vocabulary-controlled "basal" readers and phonics instruction while emphasizing the use of journal writing to develop print vocabulary.

In second grade, *Bridge* and Whole Language children were mixed in classes. The main program continued to be Whole Language, with emphasis on journal writing and reading individually selected books. No special intervention was made with the *Bridge* children.

The *Bridge* program was not used with all of the children in the *Bridge* classrooms. Two of the teachers chose only those children who were not making progress with Whole Language methods. The third teacher included all first-grade children. In order to create equivalent groups for comparison, we needed to be able to identify those first-grade children in both the *Bridge* and the Whole Language groups whose pretests indicated that they were at risk for reading failure. For this purpose, a battery of seven tests that had proven effective as predictors of first-grade reading progress was administered. (These are described under "Measures.")

At the end of the school year, we used stepwise regression analysis to determine which pre-program tests were effective in predicting *Wide Range Achievement Test* (Jastak & Wilkinson, 1984) reading scores in the Whole Language group. (Seventy-one first-grade children were available for this analysis.) Using the regression equation thus determined, we then selected all children in both groups who were predicted on the basis of pretreatment scores to achieve WRAT reading raw scores of 40 or less.[2] These children constituted the *Bridge* and Whole Language groups for comparison of the two programs.

The outcomes were assessed by testers who were unaware of the children's experimental status.

Sample

This study involved 125 first-grade children in two inner-city schools in Toronto, Canada. More than half of the children in the sample come from subsidized housing projects. About two-thirds were non-white (e.g., Afro-Canadian, Chinese, Indian, Pakistani, and Vietnamese). The sample reported here includes all of the first-grade children who were available for at least some fall tests and spring reading tests (52 boys and 43 girls) in year 1 and 30 additional grade 1 children (19 boys and 11 girls) who participated in the *Bridge* program in year 2. Overall, 71 (32 female) children were in Whole Language classrooms and 54 (17 female) in *Bridge* classrooms. (Most classrooms were "split grades," having some kindergarten or grade 2 children in addition to grade 1 children.)

Measures

Three kinds of measures are reported: *predictive measures* administered to grade 1 children in the fall of year 1 and year 2; *reading achievement measures* administered in the spring of year 1 and year 2 to both groups (grades 1 and 2); and *oral reading measures* administered in the spring of year 1 and year 2.

Predictive measures. A battery of nine tests was used. Three of the tests were classified "reading specific." These include the following:

Recognition-discrimination (Satz, Taylor, Friel, & Fletcher, 1978). This involves choosing correct matches for geometric stimuli from four choices and has also proven effective at predicting reading problems.

Visual naming speeds using the *Rapid Naming Tests* (object, colors, and letters) of Denckla and Rudel (1974). These involve naming a set of 50 pictured objects, colors, or letters as fast as possible. In each presentation, the same five items are repeated in randomized groups of five.

Phoneme analysis using the *Lindamood Auditory Conceptualization Test* (Lindamood & Lindamood, 1978). This test involves representing the number of sounds in a word using blocks. Different sounds (e.g., /s/ /c/) are represented with different colors. The specific colors used are irrelevant.

Four tests were classified as "language-related." These included:

Working memory - sentences. In this task developed by Siegel and Ryan (1989), children are presented sentences aurally with the final word missing, and have to supply the final word for each sentence. After two, three, or four sentences, the child repeats the words supplied. Three trials were given at each level (two, three, or four sentences). If the child missed all three trials at a level, the test was stopped. Total score is the number of trials on which final words were repeated correctly and in order.

Oral cloze test using Siegel and Ryan's (1989) procedures. This test involves supplying missing words for sentences presented aurally (e.g., "The pretty little ___ put on their dresses.") Twenty items were used.

Vocabulary development using the *Peabody Picture Vocabulary Test* (Dunn, 1965). In this test, the child hears a word and points to one of four pictures that represents the words.

Oral comprehension using an orally administered version of the reading comprehension section of the *Canadian Test of Basic Skills* (see "Reading Comprehension" below). Following Curtis (1980), we administered passages and questions individually as a *reading* test until each child had missed five or more of the eight questions associated with each passage. We then administered the test orally, starting with the same passage. Testing was stopped when a child missed more than four questions on orally administered passages. Oral comprehension scores include the number of items passed in both reading and oral modes.

These seven tests were administered individually to grade 1 children in the fall of year 1. Five of these tests (oral cloze, color and letter naming speed, recognition/discrimination, and *Lindamood*) were administered to the additional first-grade children in the fall of year 2. Selection of tests was based on the most effective year 1 predictors of the WRAT, as described in the Results section.

Reading achievement measures. Three standardized measures of reading achievement were used:

Word identification using the *Wide Range Achievement Test* (Reading) (Jastak & Wilkinson, 1984).

Pseudoword decoding using the Word Attack scale of the *Woodcock Reading Mastery* tests (Woodcock, 1973).

Reading comprehension using an abbreviated form (two passages per grade level) of the Reading Comprehension subscale of the *Canadian Test of Basic Skills* (King, 1982). This abbreviated reading comprehension test was administered individually in conjunction with the oral comprehension test (described previously), following procedures developed by Curtis (1980). The reading comprehension score is the number of items passed successfully in the reading mode.

Oral reading errors. In addition, year 1 reading achievement was assessed in June in both groups by an analysis of oral reading errors on one "recently read" passage and one "current" passage. Passages were selected by teachers. Thus, each child read different passages.

About half of the study children were given the miscue assessment. This included 29 of the 60 Whole Language study children and 21 of the 40 *Bridge* study children. (The reduced number of children reflected the limited number of testing staff. Students tested were determined by availability at the time of testing.)

Using categories defined by Biemiller (1970), all non-self-corrected errors were categorized as *non-response errors* (child stops reading for 5 seconds, tester supplies word); or *response-errors* (all other substitution, omission, and insertion errors). Response errors were further categorized as *errors indicating graphic strategies* if they were graphically similar substitutions (substitution begins with the same letter as the printed word, e.g., "store" for *saw*), and as *errors indicating contextual strategies* if the sentence containing a substitution, omission, or insertion error made sense up to and including the error (e.g., substituting "mistake" for *error* in this sentence). (These are not mutually exclusive categories. A child can be coded as using both graphic and contextual information on the same error.) Coding reliability was 85% agreement on a sample of 100 errors.

RESULTS

Data on the prediction of reading achievement will be presented, followed by a discussion of the comparisons between the *Bridge* and Whole Language groups at the end of grade 1 and grade 2.

Table 1
Correlations Between October Predictor Variables and May WRAT Reading, Whole Language Group

	Predictor Variables								
Grade	*Pea.* Raw	Oral Comp.	Oral Cloz.	Work. Mem.	Col. wpm	Obj. wpm	Let. wpm	Rec./ Disc.	*Lindamood*
Grade 1									
N	65	65	65	64	65	67	50	70	61
r	.49	.21	.43	.33	.40	.43	.62	.40	.45
Grade 2									
N	36	37	37	37	37	37	32	38	33
r	.28	.01	.39	.23	.43	.41	.65	.32	.56

Note. In grade 1, correlations >= .25 are significant at the .05 level. In grade 2, correlations >=.35 are significant at the .05 level.

Predicting Reading Achievement

Predicting reading performance in the Whole Language group. Correlations between test performance at the beginning of first grade and end-of-year WRAT in grades 1 and 2 are shown in Table 1. For both years, letter-naming speed and the *Lindamood* (phonological awareness) tests were consistently effective predictors. The color and object naming speed measures and the oral cloze task were consistently correlated about .40 with WRAT outcomes. The remaining measures were less consistently effective predictors.

A stepwise regression analysis with the Whole Language group indicated that the letter-naming, *Lindamood*, and oral cloze tests at the beginning of first grade accounted for 59% of WRAT variance at the end of first grade (N= 60 children for whom data on all of the measures were available). The regression equation for estimated WRAT raw scores was 2.866 + .279 (letter naming-speed) + .495 (*Lindamood* raw score) + .916 (oral cloze score).

The average WRAT score for the 60 children with all three predictor scores was 34. The 29 boys averaged 30 whereas the 31 girls averaged 38. This difference was significant (t=3.24, p<.002).

We could not obtain scores on the predictor variables for some of the children. Among those who lacked one or more predictor scores, the average WRAT raw score was 26—lower than the average WRAT score for those who completed the three predictor tests. There were no significant sex differences between the 24 boys and 12 girls who lacked one or more predictor tests. By and large, the practical implication of these results for assessment is that if a child is untestable on the letter-naming speed, *Lindamood*, or oral cloze tests, he or she may be considered "at risk." The majority of missing measures were on the letter-naming speed test–the children simply could not name the five letters used in the test.

Accuracy in identifying "at-risk" outcomes. A simple approach to identifying "at-risk" outcomes is to examine all children estimated (on the basis of the letter speed naming, *Lindamood*, and oral cloze measures) to score lower than 41 on the WRAT. (Forty-one is the lowest score in the "grade 1, end" range provided for the WRAT.) Overall, 87% of the sample scored lower than 41. Ninety-three percent were estimated to do so. Of the five "false positive" (predicted to be at risk but actually at grade level), three were estimated to score 37 or better. Of the other two, one was estimated to score 33 and the other was a recent immigrant who was untestable on all three predictor tests. For practical purposes, in this sample estimated scores below 37 were highly reliable indicators of risk, estimated scores between 37 and 40 were warnings, and estimated scores of 41 or better were indicators of no risk. (There were no "false negatives.") Although it needs to be replicated in a population at less risk, the results are evidence that this approach is useful in identifying "at-risk" beginning readers.

Construction of the comparison sample. On the basis of these findings, we could identify children in the *Bridge* groups who were expected to achieve less than 41 on the WRAT if they had been in the Whole Language group. We could similarly identify children in the Whole Language group who would have been expected to achieve less than 41 on the WRAT. (For students missing one or more predictor variable, an estimated WRAT prediction of 26 was used.) Means on predictor variables for children meeting these criteria and having end-of-year WRAT scores are shown in Table 2. There were no significant differences between the groups on the WRAT scores estimated using beginning of first-grade scores. Significance tests on the other measures are difficult to interpret because of varying percentages of children who were able to take the tests at all. For example, the letter speed of the Whole Language group is higher, but only 64% of these children knew enough letters to complete the test without three or more errors, compared to 83% of the *Bridge* group. On the oral cloze test, the *Bridge* group was higher, with about 90% of each group being able to do the test. There were no differences between the groups on the *Lindamood* test, with about 82% of each group being able to do the test.

The high percentage of children predicted to fall below 41 on WRAT who actually did fall below 41 supports the view that these children were "at risk for reading failure." *Peabody Picture Vocabulary* test scores, while not the best predictor of reading achievement in this study, provide further evidence of the underdeveloped vocabulary of these children. The selected research samples had mean *Peabody* scores of 74 (*SD* 19.7) for the Whole Language group and 77 (*SD* 16.9) for the *Bridge* group. Mean percentiles were 14.6 and 15.4, respectively. There were no significant differences between the two groups on the *Peabody* at the beginning of first grade.

End-of-grade-1 comparisons. For children estimated at the beginning of first grade to score 40 or less on the WRAT, the mean WRAT raw score for *Bridge* grade 1 group (N = 40) was 30.9 (SD = 6.5). The mean percentile score for this group was 14.7. For the Whole Language group (N = 62) the mean raw score was 27.9 (SD = 6.5). Their mean percentile score was 9.1.

Table 2
Predictor Means for* Bridge *and Whole Language Children Expected to Score Less Than 41 on the WRAT

Group	Estimated WRAT[a]	Letter Speed	Oral Cloze	*Linda-Mood*
Whole Language				
N	64.0	42.0	57.0	53.0
Mean	31.3	43.7	8.8	14.5
SD	4.8	12.8	3.8	7.2
Bridge				
N	42.0	35.0	38.0	34.0
Mean	27.9	28.3	11.7	15.2
SD	5.9	21.4	3.3	7.0

[a]Cases with missing scores estimated at 26.

Table 3
Miscue Analysis of Whole Language and* Bridge *Grade 1 Children, 1990

Group and Passage	Mean Words	Total Words	Total Errors	Non-Respon.	Graphic Errors	Contex. Errors
Passage 1						
Whole Language	84.4	1519	138	27	49	62
Bridge	81.7	1470	100	43	22	23
Passage 2						
Whole Language	106.3	1913	233	88	47	64
Bridge	114.1	2053	219	88	63	52

A t-test indicates a significant main effect for treatment (t = 2.33, p<.02).

Neither the *Woodcock* Word Attack nor the reading comprehension tests proved usable with either of these samples at the end of grade 1. Only 15% of the children were able to do any items on the *Woodcock* and only 4% were able to do any items on the comprehension test.

Miscue analysis. The samples available for miscue analysis were somewhat smaller. For these smaller sample groups, the mean estimated WRAT score for the Whole Language children was 28.6 (N = 29) whereas the mean estimated WRAT score for the *Bridge* children was 28.9 (N=21). Thus, these miscue subgroups were drawn from the lower functioning part of the sample.

The miscue analysis yielded two useful findings. First, only 62% of the Whole Language children were able to read the teacher-defined "familiar passage" at all compared to 86% of the *Bridge* group. Those "not reading" required help with 25% or more words or had clearly memorized the passage as indicated by an inability to read single words when the experimenter pointed to them. This association was significant, chi-square=6.82 (1 d.f.), p<.01. These contrasts in numbers of children able to read *at all* are consistent with differences observed on the WRAT.

Table 3 shows miscue findings *for those able to read*. Overall, the *Bridge* group made about two percent fewer errors. Following Biemiller's (1970) suggestion that both non-response errors and graphic substitution errors reflect attention to print, we combined the two error types. It appears that the *Bridge* group attended somewhat more closely to print on both passages (65 vs. 56% of errors on passage 1, 70 vs. 58% on passage 2). This association was significant on the "new" passage (chi-square = 5.60, p<.025). There are no differences in proportions of errors indicating use of contextual information.

We conclude that among children showing *some* reading ability, the *Bridge* group paid somewhat more attention to print. Paying attention to print was also associated with faster progress in the original Biemiller (1970) study of first-grade children learning to read.

Grade 2 comparisons. While the experimental effect is statistically significant on the WRAT at the end of grade 1, the magnitude of the treatment effect is not large. When the *Bridge* and Whole Language groups were followed through grade 2, more substantial effects on the WRAT appeared (raw scores of 44 vs. 36, see Table 4). A t-test showed that the treatment effect was significant at the end of second grade (t = 2.54, p<.01). Clearly, the children who had the *Bridge* program were continuing to acquire sight vocabulary at a more rapid rate, even though they were no longer using the *Bridge* curriculum.

No significant differences between groups were found on the decoding and comprehension measures, although the direction of differences favored the *Bridge* group. On the *Woodcock* Word Attack (decoding unfamiliar words), mean raw scores equaled 6.1 (*SD* 8.5) for the Whole Language group, and 7.9 (*SD* 8.9) for the *Bridge* group. On the reading comprehension test, mean

Table 4
WRAT Reading Raw Scores by Treatment Group

		Grade 1			Grade 2		
Group	*N*	*Mean*	*SD*	P'tile	*Mean*	*SD*	P'tile
Whole Language	40	28.2	5.3	8.8	37.1	12.0	9.9
Bridge	30	31.0	5.1	14.0	44.1	8.2	17.1

raw scores equaled 23.1 (*SD* 17.6) for the Whole Language group, and 27.5 (*SD* 15.3) for the *Bridge* group.

DISCUSSION

This study involved a two-year comparison of the effects of using the *Bridge* reading program in first-grade classrooms using Whole Language reading instruction. The results indicated significant differences in word identification in favor of the *Bridge* program at the end of first grade and more substantial differences by the end of second grade. No significant effects were found for decoding (as measured by the *Woodcock* Word Attack test) or for reading comprehension, although trends for both measures favored the *Bridge* group.

The basic finding that grade 1 *Bridge* instruction facilitated word-identification skill in grade 2 for these highly "at-risk" children suggests that acquiring a larger sight vocabulary in first grade helped the children profit more from the Whole Language instruction they received in second grade. This finding was an encouraging one. It was supported by the observation that a significantly larger percentage of *Bridge* children could read teacher-selected passages for miscue analysis.

In addition, we found that *Bridge* children made more errors indicating attention to print. This was consistent with Biemiller's (1970) observation that in beginning reading, graphic and non-response errors are indicators of progress in learning to read.

The lack of effects concerning decoding was not surprising. No effort was made in either program to build phonological or decoding skills. We believe that it might be profitable to do so as the children acquire an effective sight vocabulary.

It should be noted that the research design was not a true "randomized" experiment because teachers assigned themselves to the *Bridge* or Whole Language treatments. The disadvantage is that we do not know how children with teachers who preferred not to use *Bridge* would do in a *Bridge* program. On the other hand, this design meant that both programs were implemented by teachers who believed that they were serving their children well. We anticipate that should the program be implemented elsewhere, it will initially be done by teachers who believe the new program will be beneficial rather than harmful.

One could ask whether the program described here is as effective as others used with disadvantaged population, for example, DISTAR (Becker, 1977). We see the two programs as complementary in the sense that *Bridge* begins with a focus on word identification for children having difficulty with this, and is strictly an introduction to sight-word reading, whereas DISTAR begins with a focus on phonology and decoding. Our view is that focus on phonology is better introduced when the children have some sense of what reading is, and what the value of a decoding system would be.

Are we reporting an effective program for children at risk (*Bridge*) or a particularly ineffective comparison group? We do know that the levels of first-grade WRAT and *Woodcock* scores obtained in Whole Language classrooms in this study were similar to results obtained in two inner city schools in the same city by another researcher (Vandervelden, 1992). Previous research has found similar poor results for Whole Language programs applied to "at risk" innercity children (Stahl & Miller, 1989). We can only restate that beginning reading with this structured program resulted in measurably more progress both during the period when the program was taught (grade 1) and following it (grade 2) than beginning reading with Whole Language methods as used by experienced teachers.

It is possible that other structured programs (e.g., standard basal, DISTAR) would have had similar or better results on word-identification measures. However, it is worth noting that the sample selected for this research was functioning at a lower level in vocabulary knowledge of English (mean *Peabody* score of 74) than is the case in many of the studies cited by Stahl and Miller (1989) in which basal instruction was found superior to Whole Language.

We believe that a more complete focus on mastering the *Bridge* sight vocabulary for three or four months followed by a wider ranging reading (including decoding) and writing (including journals) program would be more fruitful for the majority of the disadvantaged students with whom we have worked. Along with the *Bridge* program, we would include a more intensive oral language program, including reading to children (with focus on vocabulary and comprehension), referential communication tasks, and other oral language activities. Further research is needed to evaluate such a comprehensive approach to early reading instruction for "at-risk" learners.

REFERENCES

Becker, W. (1977). Teaching reading and language to the disadvantaged–what we have learned from field research. *Harvard Educational Review*, 47, 518-544.

Biemiller, A. J. (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly, 6*, 75- 96.

Biemiller, A. J. (1977-1978). Relationships between oral reading rates for letters, words, and simple text and the development of reading achievement. *Reading Research Quarterly, 13*, 223-253.

Blachman, B. A. (1984). Relationship between rapid naming ability and language analysis skills to kindergarten and first-grade achievement. *Journal of Educational Psychology, 76*, 610-622.

Byrne, B., Freebody, P., & Gates, A. (1992). Longitudinal data on the relations of word-reading strategies to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly, 27*, 141-151.

Cantalini, M. (1987). The effects of age and gender on school readiness and school success. Unpublished doctoral dissertation, University of Toronto, Toronto, Canada.

Curtis, M. E. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72*, 656-669.

Denckla, M. B., & Rudel, R. G. (1974). Rapid automatized naming of pictured objects, colors, letters, and numbers by normal children. *Cortex, 8*, 186-202.

Dewsbury, A., Jennings, J., & Boyle, D. (1983). *Bridge reading*. Toronto, Ont.: O.I.S.E. Press.

Dolch, E. (1936). Basic sight vocabulary. Elementary School Journal, 36, 456-560. (The Dolch list is reprinted as Appendix A of Barr, R., Sadow, M., & Blachowicz's *Reading Diagnosis for Teachers*, second edition, New York: Longman, 1990.)

Dunn, L. (1965). *Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.

Edelsky, C. (1990). Whose agenda is this anyway? A response to McKenna, Robinson, and Miller. *Educational Researcher, 19*(3), 7-11.

Gleitman, L., & Rozin, P. (1972). Teaching reading by use of a syllabary. *Reading Research Quarterly, 8*, 447-483.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6-10.

Jastak, S., & Wilkinson, G.S. (1984). *Wide Range Achievement Test*. Wilmington, DE: Jastak Associates.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Juel, C., Griffith, P.L., & Gough, P.B. (1986) The acquisition of literacy: a longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*, 243-255.

King, E. (1982). *Canadian Tests of Basic Skills*. Toronto, Canada: Nelson Canada, Ltd.

Lindamood, C. H., & Lindamood, P. C. (1978). *Lindamood Auditory Conceptualization Test*, Revised Edition. Allen, Tx: DLM Teaching Associates.

Lovett, M. (1987). A developmental approach to reading disability: Accuracy and speed criteria of normal and deficient reading skill. *Child Development, 58*, 234-260.

McKenna, M. C., Robinson, R. D., & Miller, J. W. (1990). Whole language: A research agenda for the nineties. *Educational Researcher, 19*(3), 3-6.

Mounoud, P. (1986). Similarities between developmental sequences at different age periods. In I. Levin (Ed.), *Stage and structure*. New York: Ablex.

Perfetti, C. (1985). *Reading ability*. New York: Oxford University Press.

Satz, P., Taylor, H. G., Friel, J., & Fletcher, J. (1978). Some developmental and predictive precursors of reading disabilities: a six-year follow-up. In A. L. Benton & D. Pearl (Eds.), *Dyslexia: An appraisal of current knowledge*. New York: Oxford University Press.

Siegel, L. S. (in press). The development of reading. In H. W. Reese (Ed.), *Advances in child behav-*

ior and development. New York: Academic Press.

Siegel, L. S., & Ryan, E. B. (1988). Development of grammatical-sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology, 24,* 28-37.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973-980.

Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research, 59*, 87-116.

Stanovich, K. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-406.

Stanovich, K. E., Nathan, R. G., & Zolman, J. E. (1988). The developmental lag hypothesis in reading: longitudinal and matched reading level comparisons. *Child Development, 59*, 71-86.

Vandervelden, M. (1992). *Phonological recoding and phonological analytical skill in early literacy learning: A developmental approach.* Unpublished doctoral dissertation, University of Toronto.

Vellutino, F. R. (1979). *Dyslexia: Theory and research.* Cambridge, MA: MIT Press.

Vellutino, F. R. (1991) Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. *Journal of Educational Psychology, 83*, 437-443.

Woodcock, R. W. (1973). *Woodcock Reading Mastery Tests*. Circle Pines, MN: American Guidance Service.

NOTES

We wish to thank the Ontario Ministry of Education and the Ontario Institute for Studies in Education for their generous support of this project through the Transfer Grant Program. Thanks are due to the Toronto Board of Education, Principals Byron Grant and David Gladstone, and especially the teachers without whose participation the project would not have been possible. We thank our research associates, Elizabeth Morley and Margaret Lesperance. We also wish to thank Robbie Case who first suggested this project and who made many constructive suggestions on the design of the study.

ENDNOTES

[1]Two of the first three teachers transferred to other schools where one continued to use Bridge and the other taught fourth grade. The third transferred to third grade in the same school.

[2]This yields either the 47th or 19th percentile, depending on the age of the child at testing.

Correspondence should be addressed to: Andrew Biemiller, Institute of Child Study, University of Toronto, 45 Walmer Road, Toronto, Ontario, Canada M5R 2X2.