DOCUMENT RESUME

ED 384 003 CS 012 167

AUTHOR Gee, Eric J.
TITLE The Effects of a Whole Language Approach to Reading Instruction on Reading Comprehension: A Meta-Analysis.
PUB DATE Apr 95
NOTE 28p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; Elementary Education; *Instructional Effectiveness; Meta Analysis; *Reading Comprehension; *Reading Instruction; Reading Research; Research Methodology; Research Needs; *Whole Language Approach
IDENTIFIERS *Direct Instruction; Research Suggestions

ABSTRACT

A meta-analysis used Glassian techniques to compare the effectiveness of the whole-language approach to the direct-skills approach. A total of 21 studies included sufficient information to calculate effect sizes and included whole language approach as the independent variables with dependent variables relating to reading improvement. Results indicated: (1) nearly every study analyzed showed a positive effect size in the direction of the whole language approach; (2) the overall effect size demonstrated a significant difference between control and experimental group; and (3) studies employing random assignment had significantly larger effect sizes. However, the majority of the studies (85%) employed a quasi-experimental design which may severely limit the interpretation of the effectiveness of whole language. Further research should attempt to use random assignment and larger sample size to make the effectiveness of whole language more salient. (Contains 46 reference. An appendix of data is attached.) (RS)

The Effects of a Whole Language Approach to Reading Instruction on Reading Comprehension: a Meta-analysis

Eric J. Gee

Utah State University

Abstract

Whole language has become a prevalent method of reading instruction in the last ten years. Many researchers and educators consider it a more effective method than the traditional direct-skills instruction approach. For this reason, many studies have been conducted comparing the whole-language approach to the direct-skills approach.

What is lacking is a meta-analysis of published studies dealing with specifically whole-language. Stahl and Miller (1989) conducted a meta-analysis but their study included studies looking at language experience approaches. There is some doubt that these approaches are experimentally comparable to the current whole language approach.

The meta-analysis conducted ir. this study used Glassian meta-analytic techniques. The results found that the whole-language approach did result in a higher achievement scores in general than more traditional approaches. An analysis of certain moderating variables is also included.

The Effects of a Whole Language Approach to Reading Instruction on Reading Comprehension: a Meta-analysis

With public educators sitting in the proverbial hot seat more and more, it is only natural that they would cast about for methods to improve educational quality more frantically than usual. This call for newer and better methods is nowhere more evident than in reading instruction. Reading being the foundation for learning in all other disciplines, it is important that children receive the best instruction possible. One particular method that has garnered its share of praise and censure is the whole-language approach to reading instruction.

In the late 1960's, Goodman (1965) conducted a study the results of which led him to believe that current practices in reading instruction were not effective. Goodman found that children made fewer mistakes when learning words in context as opposed to learning the words from word lists. This study became the basis for the whole-language movement in reading instruction.

Traditional reading instruction has consisted of direct instruction of discrete skills as might be found when using a basal reader or phonics instruction (Klesius, Griffith, & Zielonka, 1991). Some researchers have referred to this as the "code-emphasis" approach (Vellutino, 1991). Letters and words are treated as "codes" which have corresponding sounds the children must decipher. Once the "code" is cracked, i.e., the children learn the sounds that correspond with particular letters or combinations of letters, the child is able to read. Such instruction teaches children the individual sounds of words and directly instructs them in the identification of words. Others refer to it as the "skills-based" approach (Klesius et. al. 1991; Freppon 1993). Specific skills needed for decoding the letters and words are taught in isolation from specific

learning tasks (Klesius et. al. 1991). Basal readers tend to follow a code-emphasis or skills-based approach to reading instruction (Miller & Milligan 1989; Reutzel & Cooper 1990).

Whole-language supporters believe that the acquisition of reading skills depends upon the context in which these skills are presented. Individual words are learned more easily and fluently when presented within a particular context. The word gathers its meaning from the other words around it and from the structure of the story. Emphasis is placed upon teaching reading with real books rather than the structured stories presented in basal readers. Such instruction has also been called the "meaning-emphasis" approach (Vellutino, 1991; Chall, 1989). Justifications for this approach are that it mimics the way in which children naturally acquire language (Sinatra, 1984; Goodman & Goodman 1979; Ribowsky, 1985), presents a more enjoyable environment for children to learn in (DeBoer, 1991; Patterson, 1992; Trachtenburg & Ferruggia, 1989), and introduces the children to good books thus stimulating their interest in reading (Kasten & Clarke, 1989; Patterson, 1992).

> Thus, a whole language approach considers provision of diversified language experiences and repetitive reading aloud as far more valuable preschool literacy activities than sequential lessons that teach recognition of individual letters and sounds. (Ribowsky 1985, p. 7)

Edelsky, Altwerger, & Flores (1991) justified the use of whole language based using the following assumptions: language is learned through actual use; language learning is both natural and social; and there are universals in language learning. Universals being that understanding of language arrived at through a process of hypothesis-generation. Children learn about

language by experimenting with language not by receiving direct instruction in the mechanics of language.

An example may be helpful in pointing out the differences between the two approaches to reading instruction. A teacher using a skills-based approach would specifically teach children the effect of a silent "e" on the end of a word. The teacher would explain that "bit" uses the short "i" sound while "bite" has a long "i" sound. The difference is presence of the silent "e". A teacher using a whole language approach would not point this out to the child but rather expose the child to several words with and without the silent "e" and allow the child to arrive at their own conclusion about the rule. This example is simplistic in the extreme but it serves to point out the basic differences between the two approaches to reading instruction.

Whole language supporters are quick to point out that whole language is not a canned method but rather an overall philosophy about the way children learn and the best way to teach them (Edelsky et. al., 1991). Many different techniques for teaching reading may fall under the auspices of whole language.

Naturally, the whole language approach has had its share of critics. Nicholson (1991) attempted to replicate Goodman's original study. He concluded that Goodman had overemphasized the positive effects of context. Nicholson suggested that the process by which children learn through whole language would be better called "enlightened guessing". Two other studies found the "code-emphasis" approach to be an effective means of teaching reading instruction (see Foorman, Francis, Novy, & Liberman, 1991; Byrne & Fielding-Barnsley, 1991).

Naturally, supporters of both positions vehemently support their particular view and attack the other (see Goodman, 1987). On the surface, whole language does seem to make good

sense. However, obviousness is no substitute for empirical support (see Gage, 1991). Such support has been sadly lacking in the literature on whole language (Chall, 1983). There are plenty of articles extolling the virtues of one approach over another but few that actually attempt to provide experimental data for their position. One reason may be the disdain that whole-language supporters have for traditional research methods (Reutzel & Cooper, 1990). Qualitative methods are considered more appropriate by some researchers for studying the effects of whole-language instruction (Freppon, 1991, 1993).

A literature search conducted for the purposes of this study revealed that recent years have seen a slight surge in the numbers of whole language research studies, both qualitative and quantitative. Some studies have also reviewed the literature in this area. Unfortunately, these reviews have been, by and large, extremely deficient. Generally, the reviews consist of discussions of the theoretical aspects of whole language, a narrative review of a few studies, a statement of the author's personal position, and a call for more studies. (see Waterman, 1991; Giddings, 1992; Feng, 1992).

A more useful method for reviewing the literature in the area may lie in the use of meta-analytic techniques as first described by Glass (1976). Standard literature reviews generally "talk" their way through the review. They consider the number of studies, the general trends those studies show and attempt to draw some conclusions based on their analysis of the studies. Meta-analysis presents a more quantitative and systematic way of reviewing the literature in a particular area. A meta-analysis attempts to determine how large an effect is being produced in general by the variable in question.

The effect size is an important tool in getting at this particular question. The effect size is the mean difference between the experimental group and the control group divided by the within group standard deviation. A meta-analysis generally calculates effect sizes for the studies being reviewed and then computes an average effect size. This average effect size becomes the indicator for the overall effect of the independent variable on the dependent variable. Meta-analysis summarizes not just the studies but the individual effects reported in each study.

Meta-analysis is an effective method for a variety of problems. Glass (1976) mentioned that reading instruction is a prime area for the use of meta-analysis because so often the dependent variables are standardized test scores. This makes comparisons across studies more easily interpretable. In an area like whole language, where exists controversy over the effectiveness of the method, a meta-analysis can help draw some general conclusions. While another substantive study would certainly be welcome in the area, it would seem that an attempt to make some general statements about studies in the area is long overdue.

Despite the potential usefulness of meta-analytic techniques, thus far there has only been one performed in the area of whole language. Stahl and Miller's study in 1989 was the first (and so far only) meta-analysis of whole-language studies. They concluded that, overall, there was little or no difference between whole-language approaches and more traditional methods of reading instruction. They did find some effects for lower grades, but these effects appeared to wash out by the time students reached the upper grades. However, some justifiable criticisms of the study were raised.

Schickedanz (1990) felt that their selection of studies was too narrow to justify the broad conclusion they drew. Study selection was restricted by the nature of the populations from

which the study drew its sample. In their meta-analysis, three studies dealt with third graders and two studies dealt with second-graders. The remaining fifteen studies had first-grade or kindergarten populations.

Another inclusion criterion also presents a serious blow to the validity of their study. Not only were whole language studies included in the meta-analysis but Stahl and Miller also included what they called "language experience approach" studies. This was a particular approach somewhat prevalent in the late sixties and early seventies. McGee and Lomax (1990) declared that the language experience approach is not similar enough to current whole-language programs to justify including such studies in the meta-analysis.

> "Therefore, we would argue that whole language and language experience are related in that language experience was an important precursor to whole language, but we assert the results of a study examining language experience approaches cannot be used to make inferences about whole language as Stahl and Miller do." (McGee & Lomax, 1990, p. 135)

The language experience approach was similar in some ways to whole language. It deemphasized the use of the basal reader and emphasized the use of language experiences in teaching. However, as Edelsky et. al. (1991) points out, the language experience approach was not based on any theory about language. In addition, a statement of how learning to read should take place based on its principles is not available. Language experience approaches used transcripts of a child's own talk to teach the child to read. In addition, this approach did advocate the use of some skills instruction.

There is a need for another quantitative literature review that focuses ONLY on whole language approaches. The purpose of this study will be to perform such a meta-analysis on whole-language programs and answer the following question:

What is the effect of whole-language instruction on skills related to reading comprehension when compared to more traditional types of reading instruction?

Method

Study Selection

Studies were located through the use of the ERIC and PsyLit databases. Key words were "whole-language approach" and "reading instruction' or "reading comprehension". Studies from Stahl and Miller (1990) dealing specifically with whole language were obtained where available. Reference sections from various articles on whole language were reviewed for possible studies. Social Science Citations Index was also used to locate appropriate studies.

The studies included had to use the whole-language approach as the independent variable and dependent variables had to relate to reading improvement. The studies must also make a comparison between the whole-language approach and regular classroom instruction. There was no specific criteria for age of subjects. A total of 30 studies were found which met the criteria.

In the studies selected, most of the outcome instruments were designed to measure reading achievement of some kind. Seven individual variables related to reading achievement were identified in the selected studies: reading comprehension, vocabulary skills, decoding ability (related to phonics), word/letter identification, grammar, verbal intelligence, and generation of information from the context. Forty-six percent of the reported measures employed some form of reading comprehension. Fourteen percent measured vocabulary

acquisition. Eighteen percent measured decoding ability. Seven percent considered word/letter identification and another seven percent used the Cloze procedure which provided a context for the student to use to generate the answer. One measure was a grammar measure and another an index of verbal intelligence. Measurement of these variables took the form of a variety of standardized and author-constructed tests.

Another selection criteria considered the nature of the treatment and comparison groups in the studies. Sixteen of the twenty studies compare two groups. Of those sixteen, seventy-five percent describe the treatment group as using "whole language". One study (Tumner & Nesdale, 1985), described it as the psycholinguistic approach; psycholinguistics being the basis of the whole language approach (Weaver, 1988). Taylor, Blum, & Logsdon (1986) described the treatment as "language and print rich". Anderson, Wilkinson, & Mason (1991) labeled their treatment group as "meaning rich" while Bridge, Winograd, & Haley (1983) claimed the use of predictable patterns in the treatment group. All of these were judged to contain the vital elements of the whole-language approach; mainly: relevant use of language, oral reading of the text, emphasis on gaining meaning and reading material consisting of "real" literature (Klesius et. al. 1991).

Most of the treatment groups were compared to a group using a basal reader emphasizing skills or phonics instruction. In a few cases, the comparison group was described as "traditional". This usually referred to use of a basal reader or a skills-oriented approach. Three studies described their comparison group as using a "code-emphasis approach." This referred to a skills-based approach which included teaching of phonics. By and large, the comparison groups were judged to be implementing a similar instructional approach.

Procedures

Analysis of the data was conducted using Glassian meta-analytic techniques (see Glass 1976). Such analysis collects all studies relevant to the research question, effect sizes are calculated for all dependent variables in the study, and the average effect size is then calculated and described (Bangert-Drowns, 1986). Since some studies did not include sufficient information to calculate effect sizes, only 21 studies were used in the final analysis. From the 21 studies, a total of 52 effect sizes was found.

Effect sizes were calculated by subtracting the mean of the control group from the mean of the experimental group and dividing by the control group's standard deviation. Technically, the within group standard deviation should be used. However, Glass (1978) suggested that if that particular standard deviation is not available, the control group standard deviation is a viable alternative.

Concerns have been raised about the combining of different dependent variables generally used in Glassian meta-analysis. (see Bangert-Drowns, 1986). However, since most of the measures mentioned above are integral to reading comprehension (for example, without the ability to identify letters and words, no reading will take place at all), it was judged that none were significantly different from the target construct.

After effect sizes were calculated, the relationship of these effect sizes to five moderator variables was analyzed. The reason for analyzing the moderator variables is to ascertain if the variation in the effect sizes could possible be due to other characteristics of the studies rather than the treatment variable. Those variables analyzed were: use of randomly assigned groups, sample size, subject characteristics, quality of the study, and length of the intervention.

Results and Discussion

The resulting average of all the effect sizes was .65 (see appendix). This indicates that the group mean for individuals in the whole-language condition tended to lie about two-thirds of a standard deviation above the means of the individuals in the control or comparison group. A broad interpretation of this average effect size is that whole language instruction does seem to result in higher reading achievement.

This is contrary to the conclusions of Stahl and Miller (1989) who claimed they could find no significant differences between whole-language and other types of instruction. They did claim that whole language seemed to increase scores in kindergarten classes but the effects washed out by first grade. If the language experience programs were significantly different from present day programs and not as effective as the whole language programs, it may explain why Stahl and Miller found no large differences.

The majority of the studies analyzed did not employ an experimental design. Eighty-five percent of the studies used a quasi-experimental design. Fifty-five percent employed a nonequivalent-control group design. Twenty percent of the studies used a one-group pretest-posttest design and ten percent used a static-group comparison design. Only fifteen percent of the studies randomly assigned subjects to the treatment and comparison groups. A t-test was used to compare those studies using random assignment and those which did not. The mean effect size for studies using random assignment was 1.29. The mean effect size for those studies not using random assignment was .49. The difference between these two means was found to be statistically significant ($p < .01$). This may indicate that studies using random assignment tended to get higher effect sizes than those studies not using random assignment.

Despite the limitations that must be imposed on the analysis because of the preponderance of quasi-experimental designs, there are some obvious trends in the direction of the whole-language approach. If enough studies are generated, more statistical analyses can be done relating effect sizes to this threat to internal validity (Glass, McGraw & Smith, 1981).

Length of the intervention is an important variable. If treatment time is too short, the efficacy of the treatment will be in doubt. This particular variable did not seem to hamper the results of the meta-analysis. Sixty-five percent of the studies took place over the space of one school year or two school years. The shortest intervention time reported was two months. Only two of the studies (ten percent) employed that short of a treatment interval. Most of the other studies provided ample time for implementation of the treatment. On the average, interventions took place over a period of 49 weeks or six months. A correlation between the length of the intervention and the effect sizes was .27 and not found to be statistically significant. Lack of statistical significance may have been due to the relatively small sample of effect sizes that were analyzed. A cautious interpretation may be that the longer the intervention, the more effective the program.

Subjects used in the studies varied. The characteristics of the samples were not quite as homogenous as Stahl and Miller's (1990). However, the majority of the subjects were in the primary grades. Eighty-five percent of the studies used subjects from kindergarten or first, second, and third grades. Two studies (ten percent) used fifth grade students and one study used high school-aged students.

There were too few studies to draw any real conclusions about the effectiveness of whole language for the higher grades. Eliminating the higher grades from the average of effect sizes,

an average effect size of .70 is calculated. For younger students, the whole-language approach may be overall slightly more effective than for older students. Again, such conclusions should be considered with caution since there was not an adequate sample of older students.

Characteristics of those students were also fairly homogeneous. Fifty-five percent were considered "normal" students. In other words, there was an expected range of average to good students in the sample. Twenty percent of the studies dealt specifically with low achievers and twenty percent of the samples consisted of predominantly minority groups. In one study, the character of the sample was not specified. A one-way ANOVA was run on the subject characteristics (normal, learning-disabled, minority) which did not yield any statistically significant results.

Despite the lack of significance, a look at the individual means of student type yields some interesting results. Mean for "normal" students was .59, the mean for students with some sort of learning disability was .53, and the mean effect size for minority students was 1.13, substantially higher than either of the other two means. It is possible that whole language programs may be more effective for minority groups whose overall reading achievement scores are often lower than the normal population's. However, such a result must be viewed with caution since the difference among the means was not found to be statistically significant.

Another study variable was the sample size. A correlation between sample size and effect size yielded a statistically significant correlation coefficient of -.38 ($p < .01$). One conclusion that might be drawn here is that as sample size increases, the effect of the whole-language approach decreases. Many of the studies with striking results may have achieved said results by using a smaller more specific population. Future studies more carefully conducted

with larger samples may find the results of whole language programs not as effective as originally thought.

The final variable considered was study quality. This was measured by separating those studies published in journals from those studies that were mainly presentations at conferences or documents on the ERIC database. Since journals have a more stringent criteria for acceptance of articles, this criteria was judged to be an adequate measure of study quality. A t-test found no significant differences between the effect sizes of those studies presented and the effect sizes of those studies published in journals.

Conclusions

As mentioned before, there seems to be some advantages to using the whole language approach over a more traditional, skills-based approach. Nearly every study analyzed showed a positive effect size in the direction of the whole language approach. The overall effect size demonstrated a significant difference between control and experimental groups.

However, some characteristics of the studies would seem to indicate some qualifying of this particular conclusion. The majority of the studies employed a quasi-experimental design which may severely limit the interpretation of the effectiveness of whole language. Studies employing random assignment had significant larger effect sizes. If a majority of studies employed such random assignment, the effects of a whole language program may be even more striking.

On the negative side, sample size was negatively correlated with effect size. Studies using larger, more general samples may find less striking effects from using whole language instruction. In addition, the nature of those samples may make a great deal of difference.

Despite the lack of statistical significance, the mean differences among the different sample types would seem to indicate a larger effect for minority students. This may be because minority students generally come from a literature-poor environment and whole-language's emphasis on literature produces such phenomenal changes. It may be that whole-language is more useful for that particular population and that its effects on "normal" students may not be as striking.

Another important indicator is the length of the intervention. It would appear that one school year is generally long enough to produce measurable effects. Shorter intervention periods may not provide the depth of instruction necessary to produce measurable changes in students' reading achievement.

There are some limitations to the study. First of all, a small number of studies were used in the meta-analysis. Other studies were available but insufficient data in those studies precluded the calculation of an effect size. In addition, though the effect size was not tested for statistical significance, the small number of studies MAY have resulted in lack of significance. In addition, some studies contained more dependent variables than others. The high number of effect sizes generated for one study may have skewed the results in favor of whole language. For example, Uhry & Shepherd (1993) have nine effect sizes for their study. All of which are positive.

Second, whenever a new program is introduced, there is a certain amount of enthusiasm of the part of the teacher employing it. In every study, teachers were generally implementing the treatment. Most of these teachers were already implementing the whole-language approach or had attended a workshop where they were taught whole-language methods. Such eagerness on the part of the teachers could account for the increases seen on the dependent variables. An

interesting study would be a meta-analysis of all studies where the treatment being implemented is new to the teacher.

Third, some of the moderating variables bring the internal validity of the studies selected for analysis into question. For example, the high number of quasi-experimental designs may threaten the internal validity of the analyzed studies thus threatening the accuracy of the mean effect size.

Fourth, few generalizations can be made among the higher grade levels since few studies have been conducted with that particular age group. Differences among the type of student used in the studies seem to indicate that minority groups benefit most from whole language programs, but such a statement must be tempered by the lack of statistical significance found in the one-way ANOVA of student type.

Finally, a meta-analysis can not account for qualitative data. One of the criticisms leveled against many quantitative studies of whole language is that whole language is better studied using the techniques of qualitative research; such as observation and anecdotal evidence. Many of the studies reported qualitative data along with the quantitative but, of course, this data had to be ignored in the meta-analysis (see Deboer, 1991; Freppon, 1991; Freppon, 1993; Stasko, 1991 and Stice & Bertrand, 1990).

Future studies should attempt to use random assignment and larger sample size to make the effectiveness of the whole language approach more salient. To counter the effects of teacher enthusiasm, more longitudinal studies should be conducted to see how long-lasting the effects of the whole language approach are in comparison to the more traditional approaches.

At the very least, whole language instruction is just as effective as more traditional approaches. At the most, the whole language approach may well supplant the basal reader in the years to come as a superior method of teaching reading. However, this will only happen if whole language develops a broader empirical base. If this does not happen, whole language may go the way of a number of educational "fads".

References

Anderson, R.C., Wilkinson, I.A.G., & Mason, J.M. (1991). A micro-analysis of the small group guided reading lesson: effects of an emphasis on global storymeaning. Reading Research Quarterly, 26, 417-441.

Bangert-Drowns, R.L. (1986). Review of developments in meta-analytic method. Psychological Bulletin, 99, 388-399.

Bridge, C.A., Winograd, P.N., & Haley, D. (1983). Using predictable materials vs. preprimers to teach beginning sight words. The Reading Teacher, 36, 884-891.

Byrne, B. & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. Journal of Educational Psychology, 83(4), 456-469.

Chall, J.S. (1983). Stages of Reading Development. New York: McGraw-Hill.

Chall, J.S. (1989 March). Learning to read: the great debate 20 years later--a response to "Debunking the great phonics myth." Phi Delta Kappan, 521-538.

DeBoer, J. (1991). The response of fifth grade low achievers to literature-based reading instruction through whole class and heterogenous arrangements. (Report No. CS 010 724). Michigan State University. (ERIC Document Reproduction Service No. 336 736).

Edelsky, C., Altwerger, B., & Flores, B. (1991). Whole language: what's the difference? Portsmouth, NH: Heinemann.

Eldridge, L. (1991). An experiment with a modified whole language approach in first grade classrooms. Reading Research and Instruction, 30, 21-38.

Evans, T. & Carr, T.H. (1985). Cognitive abilities, conditions of learning, and the early development of reading skill. Reading Research Quarterly, 20, 327-349.

Feng, J. (1992). Whole language approach: is it really better? (Report No. CS 011 043). Opinion paper. (ERIC Document Reproduction Service No. ED 349 548).

Foorman, B.R., Francis, D.J., Novy, D.M., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. Journal of Educational Psychology, 83(4), 456-469.

Freppon, P.A. (1991). Children's concepts of the nature and purpose of reading in different instructional settings. Journal of Reading Behavior, 23, 139-163.

Freppon, P.A. (1993 April). Making sense of reading and writing in urban classrooms; understanding at-risk children's knowledge construction in different curricula. Paper based on report presented at the annual meeting of the American Educational Research Association, Atlanta: GA. (ERIC Document Reproduction Service No. ED 361 433).

Gage, N.L. (1991 January-February). The obviousness of social and educational research results. Educational Researcher, 10-16.

Giddings, L.R. (1992). Literature-based reading instruction: an analysis. Reading Research and Instruction, 31, 18-30.

Glass, G.V. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G.V. (1978). Integrating findings: the meta-analysis of research. Review of Research in Education, 5, 351-379.

Glass, G.V., McGraw, B. & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Goodman, K.S. (1965). A linguistic study of cues and miscues in reading. Elementary English, 42, 639-643.

Goodman, K.S. (1987, August). To my professional friends in Australia and New Zealand. ARA Today: A Quarterly Newsletter From the Australia Reading Association, 6.

Goodman, Y. & Goodman, K.S. (1979). Learning to read is natural. In L.B. Resnick & P.A. Weaver (Eds.), Theory and practice of early reading. Hillsdale, NJ: Erlbaum.

Kasten, W.C. & Clarke, B.K. (1989). Reading/writing readiness for preschool and kindergarten children: a whole language approach. (Report No. PS 018 262). Sanibel, FL: Florida Educational Research and Development Council, Inc. (ERIC Document Reproduction Service No. 312 041)

Klesius, J.P., Griffith, P.L., & Zielonka, P. (1991). A whole language and traditional instruction comparison: overall effectiveness and development of the alphabetic p principle. Reading Research and Instruction, 30, 47-61.

Manning, M. (1989 November). Effects of a whole language and a skill-oriented program on the literacy development of inner city primary children. Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. 324 642).

McGee, L.M. & Lomax, R.G. (1990). On combining apples and oranges: a response to Stahl and Miller. Review of Educational Research, 60, 133-140.

Miller, J.K. & Milligan, J.L. (1989 July-August). A comparison of the whole language approach with a basal reader approach on the decoding and comprehending ability of beginning readers. Paper presented at the European Conference on Reading, Berlin, West Germany. (ERIC Document Reproduction Service No. 313 693).

Milligan, J.L. & Berg, H. (1992). The effect of whole language on the comprehending ability of first grade children. Reading Improvement, 29, 146-154.

Nadler, D.L. (1989). Teaching reading to chapter one pre-first grade students using a literature based approach. (Report No. CS 410 231). Nova University. (ERIC Reproduction Service No. 323 518).

Nicholson, T. (1991). Do children read words better in context or in lists? a classic study revisited. Journal of Educational Psychology, 83(4), 444-450.

Patterson, R.Y. (1992). Implementing an integrated approach to reading to develop critical thinking skills among a group of first graders. (Report No. CS 011 012) Nova University. (ERIC Document Reproduction Service No. 348 659).

Reutzel, D.R. & Cooper, R.B. (1990). Whole language: comparative effects on first-grade reading achievement. Journal of Educational Research, 83, 252-257.

Ribowsky, H. (1985). The effects of a code emphasis approach and a whole language approach upon emergent literacy of kindergarten children. (Report No. CS 008 397). Technical report. (ERIC Document Reproduction Service No. 269 720)

Richardson, M., Apaza, J., & Graff, D. (1991 October). Evaluation of whole language and traditional language arts instruction using a Cloze-procedure test for reading comprehension. Paper presented at the annual meeting of the Northern Rocky Mountain Educational Research Association, Jackson, WY. (ERIC Document Reproduction Service No. 339 012).

Schickedanz, J.A. (1990). The jury is still out on the effects of whole language and language experience approaches for beginning reading: a critique of Stahl and Miller's study. Review of Educational Research, 60, 127-131.

Sinatra, R. (1984 December). How a holistic language arts curriculum influenced the reading and writing proficiency of kindergarten and primary grade children. Paper presented at the annual meeting of the Eastern Regional Conference of the International Reading Association, Baltimore, MD. (ERIC Document Reproduction Service No. 264 583).

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.

Stasko, M. (1991). Increasing reading comprehension and vocabulary retention skills by using the whole-language approach. (Report No. CS 010 531). Nova University. (ERIC Reproduction Service No. 331 013)

Stice, C.R. & Bertrand, N.P. (1990). Whole language and the emergent literacy of at-risk children: a two year comparative study. (Report No. CS 010 255). Nashville, TN: Tennessee State University, Center of Excellence: Basic Skills. (ERIC Document Reproduction Service No. 324 636).

Taylor, N.E., Blum, I.H., & Logsdon, D.M. (1986). The development of written language awareness: environmental aspects and program characteristics. Reading Research Quarterly, 21, 132-149.

Trachtenburg, P. & Ferruggia, A. (1989 January). Big books from little voices: reaching high risk beginning readers. The Reading Teacher, 284-289.

Trenholm, D.S. (1992). What is the effect of traditional language teaching method versus whole language teaching on learning sentence structure? (Report No. CS 011 154). Technical report. (ERIC Document Reproduction Service No. 353 552).

Tumner, W.E. & Nesdale, A.R. (1985). Phonemic segmentation skill and beginning reading. Journal of Educational Psychology, 77, 417-427.

Uhry, J.K. & Shepherd, M.J. (1993 September). Segmentation/spelling instruction as part of a first grade reading program: effects on several measures of reading. Reading Research Quarterly, 219-233.

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.

Waterman, D.C. (1991). Whole language: why not? Contemporary Education, 62, 115-119.

Weaver, C. (1988). Reading process and practice. Portsmouth, NH: Heineman.

APPENDIX

STUDY	GRADE	N	DEPENDENT VARIABLES	EFFECT SIZE
Bridge, et. al. (1983)	1	16	Vocabulary test	1.1
Evans & Carr (19[illegible])	K-3	NA	Stanford Achievement Test Cloze Task Cell Three Passages	 -1.1 -.86
Ribowsky (1985)	K	53	Metro. Achievement Test letter recognition phoneme/grapheme	 .48 .55
Tumner & Nesdale (1985)	1	63	Peabody Picture Vocabulary Reading Assessment System real word decoding pseudoword decoding reading comprehension Phonemic Segmentation Test	 1.1 .91 1.0 .42
Taylor, et. al. (1986)	K	NA	Written Language Awareness Bohm Test of Basic Concepts 1 Bohm Test of Basic Concepts 2 Metropolitan Readiness Test	.38 .41 .50 .32
Manning (1989)	K-2	22	Stanford Achievement Test reading comprehension word study	 2.0 2.2
Miller & Milligan (1989)	1	66	Nonsense Word Test Deletion Test	.27 2.3
Nadler (1989)	K	14	Dolch Basic Sight Word Sum. Recognition of Letters Recognition of Rhyming Words	2.4 .76 1.4
Trachtenburg & Ferrugia (1989)	1	14	Comp. Test of Basic Skills	1.7
Reutzel & Cooper (1990)	1	91	Gates-McGintie Reading Test	1.2

Stice & Bertrand (1990)	1	100	Stanford Achievement Test reading comprehension Concepts About Print Retelling Scores Semantic Acceptability Corrected Miscues	 .25 .69 .85 .26 1.2
Anderson et. al. (1991)	3	149	SRA Achievement Test recall of propositions recall of story elements	 .09 .19
DeBoer (1991)	5	8	Gates-McGintie Reading Test	1.2
Eldredge (1991)	1	56	Gates-McGintie Reading Test Phonics Test	.65 .43
Klesius et. al. (1991)	1	112	Comp. Test of Basic Skills vocabulary comprehension Phonemic Awareness Test Nonsense Words	 -.41 -.11 -.80 -.72
Richardson et. al. (1991)	3	88	Cloze Procedure verbatim answers acceptable answers	 .24 .12
Stasko (1991)	5	19	Burns-Rowe Reading Inventory Academic Scores	.27 -.85
Milligan & Berg (1992)	1	165	Cloze Procedure	.24
Trenholm (1992)	9-12	60	Diagnostic English Test	.34
Freppon (1993)	2	17	Normed Measures of Reading	.31
Uhry & Shepherd (1993)	1	22	Woodcock Reading Test word attack (vowels) word attack (letters) word identification Gray Oral Reading Test Gates-McGintie Reading Test ITPA Sound Blending Subtest Roswell-Chall Blending Test Rosner's Test of Skills Block Segmentation	 1.8 2.0 1.2 .70 .46 2.0 .89 .97 .94