ADI NEWS

VOLUME 9, NUMBER 2 — THE ASSOCIATION FOR DIRECT INSTRUCTION — WINTER, 1990

ADI

Eugene Conference information inside!

See Pages 44 & 45

Direct Instruction Conferences 1990

Orlando , Florida • June 11– 15
Second Orlando Direct Instruction Conference

Lewes, Delaware • July 16—19
Sixth Atlantic Coast DI Conference

Kalamazoo, Michigan • August 6—10
Kalamazoo Conference

Eugene, Oregon • August 6—10
16th Eugene DI Conference

Salt Lake City, Utah • August 15—17
Fifth Salt Lake City DI Institute

Madison, Wisconsin • August 13—15
Wisconsin Direct Instruction Conference

Reading, Pennsylvania • August 20—22
The Third Eastern Pennsylvania Conference on DI

Seattle, Washington • August 22—24
Puget Sound DI Conference

The **Direct Instruction News** is published Fall, Winter, Spring and Summer, and is distributed by mail to members of the Association for Direct Instruction. Membership and subscription information may be found on page 41 of this newsletter. Readers are invited to submit articles for publication relating to DI. Send contributions to: The Association for Direct Instruction, P.O. Box 10252, Eugene, Oregon, 97440.

Learning to Read: The Great Debate 20 Years Later–A Response to "Debunking the Great Phonics Myth"*

by Jeanne S. Chall
Harvard University, Cambridge, Mass.

I wish to thank the editor of the *Kappan* for inviting me to respond to "Debunking the Great Phonics Myth."[1] The question of how to start children on the lifelong task of learning to read is a very old one that touches all of us – students, educators, and researchers. Increasingly, economists and legislators share in recognizing that the ability to read well is basic to our national survival. Therefore, any article that promises new insights into the ways we can best teach children to read is usually met with anticipation and excitement. Unfortunately, "Debunking the Great Phonics Myth" falls far short of its promise.

Essentially, the *Kappan* article reviews a fraction of the studies analyzed 20 years ago for *Learning to Read: The Great Debate*[2] and claims that "errors" in reporting and methodological "flaws" in some of the studies analyzed make the conclusions drawn in *The Great Debate* untenable. "Debunking . . ." Further suggests that any later research that cites *The Great Debate* is equally untrustworthy.

I should like to say at the outset that I found the article rather unreal – from its opening statement of purpose to its concluding remarks. The article begins, "My purpose in writing this article is not to enter into the century-old debate over whether an emphasis on the code or on meaning works best in teaching reading."[3] It soon engages in what most readers would say is a relentless debate *against* phonics and *for* whole-language instruction, the current meaning-emphasis approach. It asserts strongly at the start that the low state of reading achievement in the U.S. is attributable to an emphasis on phonics, yet it provides little evidence to support these assertions. A negative opinion on the 27 first-grade cooperative studies (sponsored by the U.S. office of Education) is quoted at length. Yet the conclusion of the directors of the coordinating center – that more phonics produced better results – was ignored. Moreover, the names of the directors, Guy Bond and Robert Dykstra, appear only in a long footnote.

Although the debunking in the *Kappan* article concentrates on only a part of *The Great Debate*, the reader is not told what was omitted and why. The article further assumes that current syntheses on the phonics issue (e.g., *Becoming a Nation of Readers*[4]) suffer from the same shortcomings because they acknowledge the earlier study as a source – an embarrassingly negative view of the independent scholarship of one's colleagues in research. Also confusing is the complete disregard of the new research and theory concerning the issues treated in *The Great Debate*. Of course, there is a place for persuasive writing, but when an article is published in the prestigious journal of Amercia's largest professional society in education, one expects that the norms of science and scholarship will be respected.

Overall, the *Kappan* article strikes me as irresponsible and possibly harmful. The major goal of my response, therefore, is to undo some of this damage.

I should like to make clear that, although I was puzzled by the content, tone, and style of the article and by its attitude toward science and scholarship, I approached the article as the product of an honest researcher seeking honest answers. Hence, I have tried to clear up the confusions and inaccuracies it contains with regard to *The Great Debate* and the update and also to bring together the new evidence on the issue of beginning reading and phonics. I conclude with a statement of the best available evidence on this issue – for education policy and for practice.

My response is presented in four parts and in an Appendix. Part 1 takes up the question of whether reading achievement among American children declined in the past two decades and whether teaching phonics was the cause. Part 2 deals with the critique of *The Great Debate* (1967 and 1983) in the *Kappan* article – its allegations regarding "errors" and "flaws" – and my responses. Part 3 presents the research and theory on beginning reading and phonics that have been published since 1983. Part 4 presents my position on policy and practice with regard to beginning reading and phonics – a sort of guide for the perplexed. Finally, an Appendix deals with criticisms leveled at the individual studies analyzed in *The Great Debate*, along with my responses.

Part 1, Reading Achievement and Phonics Instruction

The *Kappan* article decries the low ranking of the

* From the *Phi Delta Kappan*, Vol. 70, pp. 521-538, March, 1989. Reprinted with permission of the author, Jeanne S. Chall. Copyright 1989.

U.S. in literacy and strongly implies that phonics, which "has been emphasized in many classrooms for the past 20 years," is the cause.[5] This is misleading. First, from 1970 to 1980 the reading of 9-year-olds, as measured by the National Assessment of Educational Progress (NAEP), improved significantly, Furthermore, the greatest gains were made by groups that were previously the lowest achievers.[6]

The *Kappan* article reported accurately the increase in phonics teaching from 1970 to 1980. Helen Popp presented data on this trend as it was developing.[7] The NAEP noted that "the decade of the 1970s in particular was an era of emphasis on the 'basics' . . . "[8] But the increase in phonics instruction and the emphasis on "basics" were associated with higher, not lower, reading scores.

The period from 1970 to 1980 saw other changes that accompanied the rise in reading scores. Basal readers contained larger vocabularies,[9] made possible because children could "unlock" new words better than when a sight-word approach was followed.[10]

What about the 1980s? Here there is less certainty – regarding both student achievement and the factors related to it. Using the NAEP data again, we find that scores for 9-year-olds in 1984 leveled off when compared to scores for 9-year-olds in 1980.[11] In 1986 the scores had dropped so considerably (about half a grade level) that a committee was appointed to study the situation. Although the committee concluded that the steep decline between 1984 and 1986 came mainly from changes in test administration, changes in instruction were not ruled out as a contributing factor.[12]

Is this decline in scores real?[13] If so, is the emphasis on phonics the cause, as the *Kappan* article strongly implies? The evidence suggests a decline in phonics instruction during this period and an apparent increase in emphasis on reading comprehension, even in the first and second grades.[14]

A decline in phonics instruction during the 1980s is suggested by Linda Meyer, C. Nicholas Hastings, and Robert Linn, who found that the 1986 first-grade basal reading program of a major publisher provided less instruction in phonics than did that publisher's 1978 edition.[15] Similarly, Nancy Neill found that the manuals for teachers that accompanied the 1983 and 1985 first-grade readers devoted less space to teaching decoding and more space to teaching comprehension and multiple meanings of words than those published in the 1970s.[16]

Thus, if the decline in reading achievement in the 1980s is real, it may stem in part from *less* – rather than *more*–phonics being taught.[17] Indeed, further analysis of the 1986 NAEP data indicates that the decline among 9-year-olds was greatest among the lowest achievers, those who usually benefit most from phonics. Overall, the NAEP data from 1970 to 1986 seem to support the conclusions of *The Great Debate*.

The assertion made in "Debunking the Great Phonics Myth" that the low reading achievement of American children during the past 20 years comes from the increased teaching of phonics is not supported by the data. Indeed, when the achievement trends and the extent of phonics instruction are examined more closely, the data are stronger for the positive effects of phonics teaching, particularly on the reading scores of 9-year-olds – the youngest group assessed by the NAEP.

This does not imply that the more effective reading programs consist only of phonics. As I wrote in *The Great Debate*, in *Stages of Reading Development*,[18] and more recently in an article in *Principal*,[19] all effective reading programs expose children to a variety of activities that include a wide array of reading and writing. Indeed, it has been common practice in our Reading Laboratory at Harvard (and in many other reading laboratories as well) to have beginning readers learn the alphabetic principle, read the best of children's literature, and write stories and messages – all designed to foster a love of reading and books. Research and clinical experience indicate that, when phonics is part of beginning reading programs, children can read and enjoy more challenging books – and books of higher quality – at an earlier age.

Part 2. The Great Debate – Correcting the 'Corrections'

In this section, I will respond to the broad questions raised in the *Kappan*[19] article regarding *Learning to Read: The Great Debate*. My responses to specific points are in the Appendix.

Background and methods. If one's knowledge of *The Great Debate* came only from the *Kappan* article, that study would appear to have been a quick summary of 31 classroom experiments. In actual fact it was an extensive inquiry, based on data from many different sources and analyzed within a theoretical framework.[20]

The study was conducted during the early 1960s, when for a decade almost every basic issue in beginning reading instruction – "how to begin, when to begin, what instructional materials to use, how to organize classes for instruction – had been debated with intense heat and considerable rancor."[21] Just as today, there was a tendency to polarize issues, to overlook "the best available evidence" from the past, and to use emotionally charged language.

The Great Debate was composed of several studies designed to answer questions related to how best to teach beginning reading. It included a critical analysis of the available research from the classroom and the laboratory that compared different approaches to beginning reading, an analysis and synthesis of the correlational studies of reading achievement (e.g., the relationships between knowledge of the alphabet, phonics, and beginning reading achievement), and an analysis of the clinical research on the effects of teaching phonics to those with reading difficulties.

In order to synthesize the findings of these various studies, the different methods and approaches to beginning reading were defined and classified. Leading proponents of the various methods were interviewed, and different methods were observed in schools.

Finally, the reading textbooks and the manuals for teachers from the two reading series most widely used in the U.S. during the 1950s and early 1960s were analyzed. For historical comparisons, a content analysis of earlier editions (1920, 1930, and 1940) of the most widely used series of reading textbooks was performed.

In order to synthesize the knowledge base, including the knowledge that could be gleaned from the earliest studies, I had to take into account the weaknesses of those early studies (e.g., their less rigorous statistical tests and research designs). Since scholars referred to these earlier studies and since the studies had strengths that compensated for their weaknesses, I chose to include them. And I informed readers of their strengths and weaknesses.

I was well aware – and I stated so in *The Great Debate* – that some of the studies did not clearly report such aspects of experimental design as the ways in which the experimental and control groups were selected, the equivalence of instructional time for experimental and control groups, the time allotted to various aspects of reading, the ways in which teachers were selected, whether teachers followed the methods under study, and the like. More important, many of the studies did not specify clearly what a "method" involved; instead, they assigned labels (e.g., "phonics"), expecting their readers to understand what was meant. It was therefore necessary to set up rules for classifying the treatments. We classified a method as "look-say" if the author stated that it taught no phonics at all; "systematic phonics" if phonics was taught early and systematically; and "intrinsic phonics" if sight or thought reading was stressed with phonics introduced later and in more moderate amounts. I stated openly that:

> Many of the early studies did not use standard measures of outcomes or statistical tests of significance to determine whether the various results obtained could have been attributable to chance differences. At the same time, some of the first studies, with their small population, "homemade" tests, and simple statistical techniques, had many strengths that the later, more statistically sophisticated ones lacked, and I included them for that reason. . . . Also, they tended to "look at the learners," describing in considerable detail how the children under study approached words, what errors they made, their attitudes toward reading, how their teachers reacted. . . . Although most studies were unsatisfactory in some aspects, I assumed that all the authors were honest researchers searching for honest answers, and I looked for the grains of underlying truth to be found in each study.[12]

Theoretical base. While analyzing the often conflicting findings, it became evident that patterns could be found if the results were classified by grade level and by the reading component (e.g., comprehension, word recognition) tested. The hypothesis was that the effects of learning phonics would vary by grade level and by reading component.

Viewing the results in this way made it possible to see an advantage of a code emphasis (systematic phonics) compared to a meaning emphasis (less phonics or none), which was explained as follows: A stronger phonics approach (code emphasis) for beginning reading tends to result in higher word recognition scores early in grade 1 than does a look-say or a weaker phonics emphasis. (See the tables on pages 106 to 111 of *The Great Debate*.) Reading comprehension may be the same or lower early in grade 1. However, at the end of grade 1 or by grade 2, the stronger phonics programs produce better results in both recognition and comprehension. Thus our hypothesis was that, given time, phonics is advantageous both for word recognition and for comprehension – the ultimate goal of reading instruction. The advantage of phonics in beginning reading is in facilitating word recognition and fluency, which in turn facilitate reading comprehension, which in turn opens the world of books to children.

In order to broaden further the theoretical base from which to examine the methods of beginning reading instruction, I included in the analysis six reports by the leading investigators of the time on the diagnosis and treatment of children with reading disabilities. Two excerpts from the summary of my findings from these clinical studies may sound familiar to those who know of the most recent findings for children at risk (reported in Part 3, below).

> No matter how the readers in the six case studies had been taught initially, they all shared the same problem: extreme difficulty with *decoding* (not with comprehension) . . . There is *considerable* evidence. . . . that an initial reading method that emphasized "word," "natural," or "speeded" reading at the start and provided insufficient or inconsistent training in decoding produced *more serious* reading failures than one that emphasized the code.[13]

In addition, findings from correlational studies supported the experimental and clinical findings and are also in agreement with the current studies discussed in Part 3. We consistently see in these studies that letter knowledge has a very high association with early reading success, an association higher than mental ability and oral language facility. (For an analysis and synthesis of the correlational studies, see pages 140 to 150 in the 1967 and 1983 editions of *The Great Debate.*)

Working from a theoretical base, as well as from a synthesis of the experimental, correlational, and clinical findings of the research base, I recommended an earlier and more systematic emphasis on phonics in beginning reading programs, as well as reading for understanding. I recommended a code emphasis "only as a *beginning* reading method—a method to *start* the child on," and I did *not* recommend "ignoring reading-for-meaning practice."[24] I recommended changes to be made in basal readers to improve their content, including more literature and harder reading matter. I also recommended that library books, rather than workbooks, be used by children not working with the teacher and that writing be incorporated into the teaching of reading. (For more detail on these recommendation, see pages 307 through 313 of the 1967 and 983 editions of *The Great Debate.*)

Although it may not have been intended, the *Kappan* article's thin slice of the broad spectrum of research in *The Great Debate* tends to place it in polar opposition to literature, writing, and reading with comprehension. Teaching only phonics – and in isolation – was not a recommendation of *The Great Debate* in 1967 or in 1983. Viewed as a whole, *The Great Debate* provides a wide array of data from which to draw conclusions concerning early reading. In retrospect, it was worth the effort to develop a theory of beginning reading despite the limitations of data that were less than perfect. As Part 3 will make clear, current research conducted with more rigorous methodology continues to confirm and explicate the theory stated in 1967, for beginners in general and for children at risk – those from low-income families and those with reading and/or learning disabilities.

Empirical base, 1967 edition. The criticisms in the *Kappan* article concern alleged "errors" in reporting and alleged "flaws" in the research studies synthesized in 16 of the 31 classroom comparisons. The article reveals a lack of sensitivity to the norms of scientific inquiry in not indicating the methods of selecting for discussion the "16 of 31 reading experiments" from *The Great Debate.*

But how serious were these "errors" and "flaws"? Quite a few of the "errors" concerned the placement of findings from the word recognition subtest of the Gates Primary Reading Tests in the category of vocabulary rather than in the category of oral word recognition. Page 104 of *The Great Debate* defines the reading components and notes that such word recognition tests are classified as vocabulary tests because they are administered "silently" and required students to select the one word out of four that matches a picture, thus requiring knowledge of the word's meaning.

Many of the alleged "flaws" concerned studies that did not use tests of statistical significance. As my discussion above makes clear, I decided from the outset not to exclude such studies, if they had strong compensating features.

Some "flaws" cited include the possibly unequal provision of reading time for experimental and control groups. This is why we looked for trends that are supported by a theory. In a synthesis, one makes sense of general direction, regardless of some missing data. It is assumed that inadequacies in studies cancel one another out. And a theory that explains the findings is probably the greatest aid for understanding trends.

Other alleged "flaws" seem to result from misreadings of the study. For example, we are told that C.W. Valentine used Greek words.[25] In fact, Greek letters were used to spell English words. The *Kappan* article further states that Valentine found mixed results but that *The Great Debate* reported only positive results for phonics. A brief glance at Table 4-1A [26] indicates that the *Kappan* article was not correct.

But even more serious in an article purporting to "debunk phonics" is its failure to present any conclusions with regard to the articles that were considered "flawed." Should they have been omitted? Should the "verdicts" have been changed? If the *Kappan* article had given us a corrected interpretation regarding each of the studies, we could have determined whether a new overall conclusion was warranted. But his was not done. Instead, the *Kappan* article presented assertion after assertion about possible flaws in each study with no attempt to show whether these would change the findings from the individual studies or from all the studies combined.

The *Kappan* article further confounds the question of best methods for beginning reading instruction by concluding that whole-language instruction (undefined) plus "phonics as needed" (also undefined) is the best solution. In a scholarly magazine, one would expect to find some supporting evidence for this conclusion. But none is given. One can only presume that the author thought that a "debunking" of phonics was sufficient evidence to support any other method, as long as it did not include systematic, direct phonics. I would think that most educational researchers would require more empirical support.

Theoretical and empirical bases for the updated edition. In preparing the updated edition of *The Great Debate,* which covered the years 1966 to 1982, I found relatively more classroom and laboratory experiments than I had found for analysis in the first edition of *The Great Debate,* which covered the years 1910 to 1965. The update also included more studies of at-risk children.

The research from 1966 to 1982 had been conducted in a wide variety of setting by people from various disciplines. Yet the *Kappan* article attempts to dismiss these studies.

For example, the 27 cooperative first-grade studies sponsored by the U.S. Office of Education (USOE) are rejected by quoting the negative opinion of Russell Stauffer,[27] editor of *Reading Teacher,* without presenting the findings of Guy Bond and Robert Dykstra, directors of the center that coordinated the 27 studies. The 27 USOE cooperative reading studies compared various methods of instruction in the first grade (many were continued in the second grade and somewhat fewer in the third grade) and found that basal reading programs supplemented by separated phonics programs produced better results in reading at the end of the first grade.[28] There was some loss in the advantage of this approach at the end of the second grade and, for the few studies that continued to the end of third grade, few if any differences among instructional approaches were reported.

A reanalysis of the data by Dykstra (also omitted in the *Kappan* article) found an advantage for code-emphasis methods (stronger phonics) through the second grade.[29] In later report Dykstra wrote:

> We can summarize the results of sixty years of research dealing with beginning reading instruction by stating that early systematic instruction in phonics provides the child with the skills necessary to become an independent reader at an earlier age than is likely if phonics instruction is delayed and less systematic. As a consequence of his early success in "learning to read," the child can more quickly go about the job of "reading to learn."[30]

One way to judge the merit of a research synthesis is to compare it to subsequent syntheses. The *Kappan* article does so by citing, at great length, the results of Reginald Corders synthesis, which concluded that no conclusion could be drawn concerning the value of systematic phonics in beginning reading.[31] It is significant that Paul Diederich's conclusion regarding phonics from his synthesis for ERIC[32] (based on the Corder Report, on *Learning to Read: The Great Debate,* and on other studies) was not discussed in the *Kappan* article. In it, Diederich stated, "One of the few conclusions of reading research in which we can have a high degree of confidence is that earlier and more systematic instruction in phonics is essential."[33] Although this conclusion was not included in the *Kappan* article, Diederich is cited in footnote 13, from which we learn that he supported Corder's use of 22 doctoral candidates as research reviewers.

An alleged "error" in the update is that the Follow Through studies did not compare methods that included phonics, although I had noted that they did. I checked this again by telephoning Jane Stallings, the principal investigator for the study.[34] She confirmed in December 1988 that both of the "direct instruction" model groups did, in fact, use phonics. In one, the children used the Sullivan programmed reading system and in the other, the Distar system.* These phonics programs produced higher achievement in reading (on tests of word recognition, vocabulary, and comprehension) at the end of grades 1 and 3 than those programs that used less phonics. Thus the assertion that phonics was not a factor in the Follow Through studies and in the improved results of those using phonics is unfounded.

The *Kappan* article expressed concern about the extent and quality of the more recent experimental research reviewed in the updated edition of *The Great Debate.* If the amount and quality of the research is still inadequate, what is to be done? Are we to walk away from the existing research and rely, instead, on our intuition and powers of persuasion? Should we wait for the "well-conceived and carefully executed reading research that investigates a wide spectrum of potentially effective and practical instructional approaches" that the author of the *Kappan* article recommends?[35] If such studies can be realized, why have they not been conducted? Why, after similar recommendations by Corder in 1971, could so few researchers live up to his high standards? The truth is that it is extremely difficult to do experimental comparisons of classroom instruction. Therefore, one does not dismiss what one has but makes the best of it.

The experimental conditions compared in the update were different from those in the first edition of *The Great Debate:*

> Beginning with the 1970s, the importance of phonics teaching seemed to have become generally accepted, and the research question turned to what kind of phonics was the more effective. Among the characteristics differentiating the phonics programs are whether letter-sound relations are taught directly or inferred from words, whether instruction is given in blending the separate letter sounds [direct-synthetic phonics], or whether phonics elements are analyzed from larger units [indirect-analytic].[36]

---

* *Editors note.* These Follow Through Programs were the University of Oregon Model (using DISTAR) and the University of Kansas Model (using McGraw-Hill Sullivan). It should be noted that the Kansas model added small group instruction in sounds and blending (The Phonics Primer) to Sullivan (which gave inadequate instructions for teaching phonics). The program they used was developed under my supervision for a Title III project in Urbana, Illinois, and was based on Engelmann's phonic system for DISTAR but sequenced to fit the McGraw-Hill materials. My daughter Jill taught this program at our project as a high school senior and took it with her to Kansas where she worked with my good friend, Don Bushell, with the Kansas Follow Through Model.

Since both categories can fall within systematic phonics, one might expect to find fewer clear advantages in these classroom comparisons. Yet, on the whole, the *Kappan* article continued to refer to the comparisons as "more" or "less" phonics, rather than as comparisons of different kinds of phonics programs – direct-synthetic versus indirect-analytic as the update made clear.

The update of *The Great Debate* included correlational studies and theoretical issues in beginning reading that were also ignored in the *Kappan* article. [37] During the 15-year period following the publication of the first edition of *The Great Debate*, many studies continued to find high correlations of alphabet knowledge (in kindergarten and in the beginning of grade 1) with reading achievement at the end of grade 1.[38] For at-risk children, Jeannette Jansky and Katrina de Hirsch found early alphabet knowledge to be the best predictor of reading achievement in grade 2.[39]

The *Kappan* article did not address the theory included in the updated edition of *The Great Debate*, as it had ignored the theory in the first edition. and yet, if something works or does not work, shouldn't we want to know why? Thus I must register my disappointment that, while alleged methodological "flaws" (which do not exist in reality) are high-lighted and detailed, the basic research and theory that support the conclusions in the updated edition of *The Great Debate* are ignored.

From 1966 to 1982 there was a dramatic increase in basic research on reading.[40] Robert Calfee and Priscilla Drum pulled some of these studies together with earlier studies and concluded:

> In summary, the evidence from Gates and Backus (1923) through Chall (1967) and up to the present is fairly consistent – students are more likely to acquire decoding skills, if the instructional program provides time and varied opportunities to acquire these skills. Some students will learn the principles on their own, but many will not in the absence of systematic training.[41]

Part 3. Current Research Evidence on Phonics

In this section I present some of the findings from the research on beginning reading that has been published since the second edition of *The Great Debate* – from 1983 to 1988. I present these findings because, according to the norms of science, the confirmation or disconfirmation of independent scholars establishes the validity of a piece of research or theory. Shortcomings in research design are to be avoided where possible, but in the human sciences they can rarely be avoided altogether. The essence of scholarship, then, is in creating a theory, an explanation, that best fits the research. Synthesizing research, to a great extent, is creating order out of chaos. The research presented here is relevant to beginning reading and comes from cognitive psychology, psycholinguistics, developmental psychology, neurology, and the study of learning disabilities.

Cognitive psychology. Charles Perfetti, who has studies reading as a cognitive process, states in *Reading Ability* that the foundation of comprehension is accurate word recognition, which is attained through careful decoding and much practice over time.[42] Therefore, any program that results in better word recognition at the early stages – and especially with disabled readers – provides an advantage to students, and this advantage is bound to pay off over time.[43] With regard to the influence of word recognition and decoding on reading comprehension, Perfetti writes, "Evidence suggests that . . . word-processing efficiency leads to better comprehension, rather than merely being a by-product of comprehension."[44]

Keith Stanovich, who has conducted numerous studies on how reading is acquired among general and clinical populations, notes how inadequate word recognition can lead to inadequate reading comprehension:

> Slow, capacity-draining word-recognition processes required cognitive resources that should be allocated to higher-level processes of text integration and comprehension. Thus, reading for meaning is hindered, unrewarding reading experiences multiply, and practice is avoided or merely tolerated without real cognitive involvement.[45]

Marcel Just and Patricia Carpenter, who have done considerable research on eye movements, note that reading is a "letter-mediated" rather than a "whole-word-mediated" process.[46] With regard to beginning reading and decoding processes, they write:

> The accuracy of word decoding without context is correlated with measures of reading skill in which the text is present. Such correlations are also consistent with the hypothesis that word-decoding skill is sometimes a bottleneck in attaining fluency.[47]

Thus decoding or facility with phonics is viewed by these researchers and other cognitive psychologists as a necessary step in the acquisition of reading comprehension and other higher-level reading processes.

Similarly, Linnea Ehri, who has been investigating the shift from prereading to (read) reading among preschool children, concludes that the learning of letter-sound clues is important because they are required for storing words in memory and accessing words from memory.[48]

Lynette Bradley and Peter Bryant report a causal connection between phono-logical awareness and learning to read.[49] And Stanovich reports that "phonological awareness measures administered in kindergarten or earlier are superior to I.Q. tests as predictors of future reading achievement."[50]

Recent studies, one from New Zealand and one from Australia, present additional evidence that confirms the importance of word recognition and decoding in beginning reading. Tom Nicholson of New Zealand and his colleagues found that "poor readers use context to help with reading, whereas good readers who are skilled at decoding do not need to do so . . . . So the results of this study support the idea that learning to read involves learning how to decipher."[51]

Peter Freebody and Brian Byrne of Australia investigated the word-reading strategies of second- and third-graders and found theat 84% of the variation in comprehension achievement was accounted for by measures of phonemic awareness and decoding strategies.[52] Their major finding, which confirms the conclusion from *The Great Debate* that decoding has a delayed effect on comprehension, was that whole-word readers (termed Chinese readers) did better in comprehension in the second grade than did phonetic readers (termed Phoenician readers); that is, they read faster and struggled less. However, in the third grade, the Phoenician readers significantly outperformed the Chinese readers.

> The authors suggest that word-specific associations may serve a student adequately up to about second grade, but that failure to acquire and use efficient decoding skills will begin to take a toll on reading comprehension by grade 3. In contrast, Phoenicians may be hindered in comprehension performance in the early years, but begin to improve comparatively as they progress through school.[53]

Psycholinguistics. Ingvar Lundberg and his associates in Sweden found that phonological awareness was the single most powerful predictor of future reading and spelling skills in a group of children tested at the end of their kindergarten year.[54]

This sensitivity to the sounds of words is generally acknowledged to be causal as well as reciprocal. Segmenting and blending phonemes (separating the sounds in words and putting them together again) are considered to be the sine qua non of early reading development.[55] Moreover, when students are trained through tasks designed to develop their awareness of phonemes, their reading achievement improves.[56]

Mary Ann Evans and Thomas Carr evaluated two programs in 20 first-grade classrooms. Half had "traditional" teacher-directed instruction, using basal readers with phonics practice and application in reading connected text. The other half were taught by an individualized language-experience method. The teacher-directed phonics classrooms scored higher on year-end tests of reading achievement, including comprehension. Moreover, the language-experience classrooms did not achieve higher scores in oral language measure, although they engaged in more oral language activities.[57]

A meta-analysis by Steven Stahl and Patricia Miller, comparing whole-language and language-experience instruction with instruction using basal readers, produced essentially the same findings.[58] Only the children in kindergarten seemed to benefit more from the whole-language/language-experience approaches. In first grade, although both approaches tended to produce similar effects, children in those programs that had stronger instruction in phonics tended to score higher. The theory was that it was important in kindergarten to concentrate on the functions of written language and that it was important in first grade to concentrate on the connection between letters and sounds in spoken words.[59]

Study of reading disabilities. Students "at risk" of reading failure have long been thought to be deficient in phonological processing. According to Isabelle Liberman, "The results of research have, I think . . .justified our assumption, . . . providing evidence that deficits of phonological processing do, indeed, underlie many of the difficulties that poor readers and spellers have."[60]

Poor readers of all ages and in many different countries have difficulty with "segmental analysis of speech/the apprehension of the phonological structure of words."[61] Furthermore, dyslexic students are often unaware of how the written symbols map onto speech.[62] But these students *can* be trained to segment and blend.[63]

Research syntheses. Several research syntheses have been published since 1983. Notable among these is *Becoming a Nation of Readers*.[64] Produced by the Commission on Reading – 10 scholars and teachers with long experience and expertise in reading and related disciplines appointed by the National Academy of Education (NAE) and sponsored jointly by the NAE and the National Institute of Education – the report concluded the following:

> What does research indicate about the effectiveness of phonics instruction? Classroom research shows that on average, children who are taught phonics get off to a better start in learning to read than children who are not taught phonics[65]

The Roots of Phonics, by Miriam Balmuth, presents a historical analysis of English phonics and of the uses of phonics for hundreds of years in reading instruction.[66] More recently, Dina Feitelson presents a cross-national analysis of the effectiveness of phonics in teaching beginning reading.[67]

The most recent and most complete synthesis to date was done by Marilyn Adams for the Reading Research and Education Center at the University of

Illinois.[68] Adams notes:

> Perhaps the most influential arguments for teaching phonics are based on studies comparing the relative effectiveness of different approaches to teaching beginning reading. Collectively, these studies suggest, with impressive consistency, that programs including systematic instruction on letter-to-sound correspondences lead to higher achievement in both word recognition and spelling, at least in the early grades, and especially for slower or economically disadvantaged students.[69]

About the nature of the research, she write:

> Some of these data were collected in the realistic but, therefore, complex contexts of real classrooms. There were collected through controlled but, therefore, artificial laboratory studies. because both sorts of investigations have their shortcomings, the data do not prove the point. Yet, they do make it very difficult to dismiss.[70]

About the kind of phonics that is most effective, she write:

> One single class of instructional methods resulted in gains significantly larger than any of the other. This class of methods consisted in teaching students about letters and letter sounds, first separately and then blended together. It consisted, in short, of explicit, systematic phonics.[71]

She concludes:

> The major conclusions of the program comparison studies are based on masses of data, gathered through formal experimental procedures, and scrutinized through relatively sophisticated statistical techniques. Yet they are – point for point – virtually identical to those at which Jeanne Chall had arrived on the basis of her classroom observations and interpretative reviews of the literature.[72]

She offers a cautionary note:

> The caveat is more subtle. We know that there are enormous differences in the outcomes of any program depending upon the particular schools, teachers, children, and implementation vagaries involved. Yet, we have also seen that, despite all of these very real and significant differences, there seems to be something about that very broad class of instruction known as phonics that is of general, substantive, and lasting value.[73]

Thus independent research and theory since 1983, as well as syntheses of research have confirmed the conclusions of *The Great Debate* and of the 1983 update.

Part 4: Summary and Recommendations

I had two purposes in writing this response: First, to correct the errors and misconceptions contained in the *Kappan* article of November 1988 with regard to the credibility of *Learning to Read: The Great Debate;* second, to present an update of the research relevant to *The Great Debate.* It is fitting that this 20-year update appears in the *Kappan.* In 1977 a 10-year update, *Reading 1967–1977: A Decade of Change and Promise,* also appeared in a Phi Delta Kappa publication, Fastback 97.[74]

What does the research and theory on phonics and beginning reading say after 20 years? Overall, the recent research evidence supports the earlier finding reported in *The Great Debate.* The evidence of the last 20 years from the classroom, the clinic, and the laboratory confirms the earlier research. Furthermore, the findings from both the earlier and the later research are supported by a growing body of theory and basic research in cognitive psychology, psycholinguistics, human development, the study of learning disabilities, and the study of the reading process.

Why, then, with so much supporting evidence is there still reluctance to accept these findings – findings that have been confirmed again and again by independent scholars and in most reviews and syntheses of the research? For an answer we should perhaps look to factors other than scientific evidence.

Joanna Williams has pointed to the theoretical attractiveness of a meaning-emphasis as opposed to a code-emphasis method of teaching beginning reading.[75] A meaning emphasis (viewing beginning reading as comprehension), she noted, is seen by many as "sophisticated," while a code emphasis, or decoding, is seen as "simple-minded."

Stanovich also considers nonscientific factors. He notes:

> [U]ntil recently the negative attitude toward word recognition was so strong that an investigator who chose to study a variable related to word recognition was often accused of denying that the goal of reading was comprehension. It has been common for articles on reading education to be cited and to become well known more because of their polemical content and writing style than because of their scientific merit.[76]

In the first edition of *The Great Debate,* I noted that those professing the value of progressive and child-centered education tended to prefer a meaning emphasis, which they viewed as more open, more natural, and more self-directed. A code emphasis, on the other hand, has been associated with traditional education and "dull drill." Today, whole-language instruction – the current meaning-emphasis approach to beginning reading – tends to be associated with a natural, "developmental," and open view. It is further assumed by many that "open," "natural" reading programs that do not teach directly lead to greater cognitive development and to greater love of reading and learning, although there is little evidence to support these claims. Indeed, the limited evidence seems to indicate the opposite, particularly for children in the early grades and for those at risk.[77]

In actual practice, systematic teaching of phonics has existed within varied philosophies of education and of reading instruction – both those with greater structure and those with greater openness. And until research brings hard evidence to refute it, systematic phonics can claim to have produced at least as many children with a lifetime love of reading as a meaning-emphasis approach. Indeed, direct, systematic phonics, since it seems to produce better readers and fewer reading problems, has probably produced *more* lifetime readers than meaning-emphasis approaches.

Thus it is counter to research evidence and to the experience of practitioners for the *Kappan* article to imply that programs that include phonics do not give proper attention to reading for understanding. Consider the following statement from the *Kappan* article: "What is important is not knowledge of phonics per se – but rather, the ability to read and understand connected text."[78] No one can disagree with this statement. Yet it can mislead if one infers from it that those who teach phonics are limited in their perception of what reading is and how it is best taught – that they teach phonics only for its own sake and nothing more.

The most superficial reading of *The Great Debate* and a skimming of any book on the history of reading instruction in America will reveal that every beginning reading method in history has had reading for understanding as its goal. What the debates have been about is whether or not an early emphasis on word recognition and decoding helps in the acquisition of reading for understanding. The scientific evidence from the early 1900s to the present indicated that it does. such an emphasis contributes to better word recognition, which in turn leads to better reading comprehension. The evidence supports this trend through grade 3, and few studies have followed up their subjects beyond grade 3.

But if the advantage is to be maintained beyond the early grades, we must also consider how reading is taught and practiced in the later grades. I quote from the first edition of *The Great Debate:*

> Whether an initial code emphasis keeps its advantage in the middle and upper elementary grades, and later, depends on how reading is taught in these grades; how much the reading program stresses language and vocabulary growth and provides sufficiently challenging reading materials. If the reading programs are not challenging enough in these respects, the early advantages will probably be dissipated.[79]

It is this hypothesis that prompted me to write *Stages of Reading Development*, in which I present – and support with research and theoretical evidence – a scheme of stages of reading, from prereading to mature, skilled reading.[80] These stages reflect the changes in reading as it develops (in the reading tasks, in the different skills and abilities required of the reader, and in the different uses to which reading is put). The early stages (stages 0, 1, and 2) are characterized by learning the alphabetic principle and the fluent reading of familiar texts – learning designed to bring a child's reading "up to" his or her more advanced skills and abilities with oral language.[81] In the later stages, which usually start in grade 4, when the child reads increasingly less-familiar texts, the reading task becomes more conceptual and cognitive. It requires knowledge of more abstract, technical, and specialized words and more advanced thinking to understand the meaning of what is read.

It can be seen, then, that basic word recognition and decoding are important not only for early progress in reading, but for the later development of reading comprehension and metacomprehension.

There are, of course, certain situations in which phonics does not get the results that the scientific evidence suggests it should. Phonics is often taught incorrectly by teachers who themselves have not learned phonics or who have not received instruction in teaching it in their college classes. some teachers may inadvertently overdo the teaching of phonics, leaving little time for the reading of stories and other connected texts; others may not really teach phonics at all but rely on independent practice in workbooks. These questionable practices are not exclusive to phonics instruction. One can make similar observations about the teaching of comprehension. In many classrooms, comprehension is not "taught," only practiced in workbooks, In other classrooms, it is so overtaught in discussions and exercises that the students are given little time to read for understanding and pleasure. Every procedure and every good purpose can be, is, and has been overdone or done incorrectly.

Currently, the anti-phonics movement has taken unto itself a pro-literature, pro-writing, and pro-thinking stance, as if those who teach phonics and decoding are opposed to these obviously excellent aims. And yet the history of reading instruction teaches us that literature, writing, and thinking are not exclusive properties of any one approach to beginning reading.[82]

Indeed, the change in the early 1970s to an earlier and more systematic teaching of phonics in basal readers brought with it enlarged reading vocabularies that made possible the earlier use of better, more mature literature.

The same is true for writing. A code emphasis leads to earlier – rather than later – writing. Those children who know the letters of the alphabet write earlier.[83] Also, early readers who know phonics use it for writing and for reading.[84]

Are the main differences, then, between a whole-language approach and those approaches that include phonics to be understood only in terms of rhetoric and misconceptions? No. There seem to be important theoretical differences, as well. There is the greater resemblance of the "whole-language" approach to the "whole-word," "sight," and "thought" approaches used from the 1920s through the late 1960s. We need more research, and I would settle for less than ideal research proposed in the *Kappan* article. I would settle for the more feasible, garden-variety studies that, taken with other studies, can help us find trends. These studies will also help define whole language, since definitions of whole language seem to vary considerably.

It is not uncommon for some educators to hold erroneous views about those who teach phonics. They tend to believe that, if one teaches phonics, one cannot be concerned with the cognitive, meaningful, creative, and joyful aspects of literacy. The author of the *Kappan* article appears to hold this view, and she proposes to replace the systematic teaching of phonics with a whole-language approach. Yet we are offered no definition of her preferred approach. One can well understand that it is difficult to define, because its proponents seem to define it in so many different ways.

Some stress the openness of whole language and its informality as compared to more structured programs; some stress its combination of speaking, listening, reading, and writing; others point to the use of "real books" rather than reading textbooks; ;and still other stress that skills (including phonics) should not be taught directly but acquired from more natural reading and writing activities.[85]

Another aspect of beginning reading programs favored by the *Kappan* author and by other whole-language proponents is that phonics be taught only "as needed." In taking such a position one must confront another, more fundamental controversy in American education, a controversy that crosses curricular areas and levels of schooling. The issue is whether direct or indirect instruction yields better student achievement. Put another way, should teachers teach directly only those who cannot learn something on their own? If the policy is to teach phonics "only as needed," then we must provide the teachers with the skills and tools to know which children "need" direct phonics instruction. Surely a test of phonic knowledge could help. Yet the author of the *Kappan* article proposes that the word-analysis (phonics) subtests be dropped from standardized reading tests, substituting in their place greater and deeper measures of reading comprehension.[86] Thus teachers could evaluate only aspects of reading comprehension, even for students whose low scores in comprehension may stem from their difficulties in word recognition and phonics.

To say that teachers should teach phonics only as needed is to put a greater burden of responsibility on teachers and children than theory, research, and practice support. And it puts at even greater risk those children who need the instruction most – low-income , minority, and learning-disabled children.

I wish to make clear that I do not take the position that all is well with the reading of American children or with the instruction they receive. There is much to be concerned about. But the solution does not lie in throwing out approaches to beginning reading that are supported by research and theory and by the results of national assessments. Research evidence does not support meaning-emphasis approaches for beginning reading, even though the former are coached in a rhetoric of warmth, openness, and great promise.

In education as in national and world affairs, history teaches us that in times of desperation we have a tendency to look for global, charismatic, single solutions to very serious problems. Only after these fail – often at great cost – are we prepared to look for solutions that are more firmly based in reality. A code emphasis for beginning reading (phonics, decoding, and word recognition) is one way to help us improve the reading of our children and to help prevent reading problems, particularly among children at risk of failure. Such an approach will not cure all learning problems among all students. But research evidence, theory, and practice all show that direct instruction in phonics improves reading achievement significantly. If we add what we know from other research on reading – the importance of early exposure to print and books, of reading to children and exposing them to many books of literature and information, of using instructional materials that are not too easy or too hard, of providing instruction in vocabulary and comprehension as students' reading develops, and of providing practice in writing and readings in all areas of the curriculum – we can improve the reading achievement of our students significantly.

One must conclude that the efforts of Marie Carbo to "debunk" the "great phonics myth" have failed. The many claims she makes regarding the shortcomings of the research on the phonics issue seem to be more characteristic of her own analyses than of those of the researchers she criticizes.

Many scholarly publications guard against such misrepresentations by employing qualified readers as peer reviewers to critique articles before they are published, particularly in areas of controversy and of social importance. Such cautions are necessary to protect children – and the educational community that serves them.

Thus, although Marie Carbo tried to show that the value of teaching phonics is a myth, the research evidence from 1910 to 1988 show that the *real* "myth" is that children can learn to read English text without knowing or learning anything about phonics or letter-sound correspondences. Marie Carbo seems to believe strongly in such a myth. But, of course, she is wrong.

Appendix

Responses to Critical Comments on Specific Studies (1967 Edition)

Because of space limitations, I will respond to half of the 16 studies criticized by Marie Carbo. These are evenly distributed across the four types of concerns (e.g., "data reported incorrectly") noted in her *Kappan* article. I will respond to the criticisms study by study. For details of Carbo's criticisms, readers are advised to see "Debunking the Great Phonics Myth," *Phi Delta Kappan*, November 1988, especially pages 229-36.

Data Reported Incorrectly

Wohleber (1953)[87]

Carbo: The Great Debate "incorrectly" reported that the "SP [systematic phonics] group performed significantly better than the IP [intrinsic phonics] group in comprehension and vocabulary in the first grade" and that "word recognition" and not "vocabulary" was tested.[88]

Response: The scores of the SP group at the end of first grade were ahead (significantly) of the IP group on the word recognition subtest (a test of vocabulary) and on the sentence subtest (a test of reading comprehension) of the Gates Primary Reading Test; the scores of the SP group were also ahead on the paragraph reading subtest (a test of reading comprehension), although not significantly.

By the end of grade 2 and grade 3, the SP group was ahead (significantly) on the word recognition (a test of reading comprehension) subtests of the Gates Advanced Primary Test.

Thus, Carbo is mistaken on each point. Moreover, the Gates word recognition subtest *is* a vocabulary test, not an oral word recognition test, as the *Kappan* article claims.[89]

Edward (1964)[90]

Carbo: "Mary Edward did not find that comprehension scores were significantly higher for the SP group, as Chall claimed. No such data were included in Edward's study. For complex comprehension abilities, Edward found 'no significant differences. . . between the two groups.' "[91]

Response: Edward *did find* that comprehension was significantly higher for the SP group. A direct quote from Edward reads:

> The experimental boys and girls [those using systematic phonics] read faster and more accurately, had larger vocabularies, comprehended better, and were more able to retain factual information than the control boys and girls. However, when the more complex comprehension abilities of organization and appreciation were examined, no significant differences were found between the two groups.[92]

With the regard to the "complex comprehension abilities," see my response to the questions raised about the Agnew study, below.

Findings That were Excluded

Morgan and Light (1963)[93]

Carbo: "Chall omitted a highly significant result on vocabulary (p<.001) that favored the IP method. However, she listed a finding in comprehension that did not approach significance but favored the SP method."[94]

Response: Table 4-2C, on page 110 of *The Great Debate*, provides separate results for two comprehension tests, but only one result for vocabulary. Actually, there were two vocabulary measures, and they both favored the IP methods. But note that the table shows an advantage in vocabulary for IP, which the *Kappan* article said was omitted! Two comprehension scores were reported because they conflicted, but only one result was reported for the two vocabulary measures because they agreed. It is accepted practice in meta-analyses to report only a single effect size per measure per study in order to avoid giving too much weight to studies that use multiple measures. Even though *The Great Debate* predated the wide use of meta-analysis, the procedure was justified. Furthermore, having two outcomes favoring IP methods under vocabulary would not change the major conclusion of the study.

Agnew (1935 and 1939)[95]

Carbo: Many complaints were expressed with reference to Donald Agnew's study of programs in Raleigh and Durham. For details, readers should consult Carbo's article.[96]

Response: We treated the 1935 and 1939 studies of Agnew as one study, as Agnew treated them. He collected data in Raleigh (1935) and in Durham (1939), but the two data collections were part of *one* planned study. This is reinforced by his summary. The distinctions in the study are between Raleigh, in which relatively little phonics was taught, and Durham, in which a lot of phonics was taught. In Chapter 5 of his 1939 monograph, Agnew writes about the relative effects of phonics instruction *within* Raleigh, which overall taught little phonics. In Chapter 6 he contrasts Raleigh with Durham, thereby comparing a modest amount of phonics with a large amount of phonics.[97]

In the 1939 experiment, significant differences favoring the SP group were found on word recognition and on two of the four comprehension measures.[98] On the other two comprehension measures, the two groups were virtually equal. Agnew interpreted these results as favoring phonics for developing "independence in word recognition, ability to work out the sounds of new words, efficiency in in word pronunciation, accuracy in oral reading, certain abilities in silent reading, and the ability to recognize a large vocabulary of written words, " but not necessarily for interest and "social reading."[99]

It is interesting that the *Kappan* article fails, as others have before, to capture the significance of a stronger phonics program as compared to a weaker phonics program. By third grade, the Agnew study shows definite advantages for the stronger phonics group – not only in phonic ability, word recognition (pronunciation), and accuracy of oral reading, but also in vocabulary, literal comprehension (directions and details), and eye-voice span. No differences were found in comprehension questions designed to measure predicting outcomes, reading for general significance, and speed of silent reading. The "no advantage" results were in the so-called higher-level cognitive aspects of reading. This is essentially the course of development predicted for a code emphasis by our developmental theory.[100]

Classification Problems

Gates (1927)[101]

Carbo: In this study, "the method that Chall classified as IP (intrinsic phonics) apparently contained no phonics at all. Compared to the phonics-taught youngsters, those taught *without phonics* scored consistently higher on all the reading tests included in the study. However, since Gates' no-phonics method was classified by Chall as a type of phonics, Gates' results were tabulated in *The Great Debate* in support of phonics. . . No tests of significance were included in Gates' study. However, 15 sets of reading scores were listed. Eleven favored the no-phonics group (classified by Chall as intrinsic phonics), two were equal, and two favored the SP group (out of six phonics tests)."[102]

Response: Gates' "intrinsic phonics" group – the experimental treatment – was, according to Gates, a type of phonics treatment, not a "no-phonics" treatment as Carbo insists. As Gates stated, "In designing the studies thus far undertaken, it was decided to attempt to compare with phonetic methods not mere unguided, 'natural,' or 'look and say' learning of reading since the inadequacies of this type of learning are strongly suggested if not fully demonstrated in various earlier investigations. Instead, improvised forms of what may be called 'intrinsic' methods have been utilized for comparison with phonetic training."[103]

Unequal Groups

Garrison and Heard (1931)[104]

Carbo: S.C. Garrison and M.T. Heard "noted that the IP and SP groups in their study differed substantially in intelligence. In the SP group, the ratio of bright to dull students was 39:18 (approximately 2:1); in the IP group the ratio was 28:26 (approximately 1:1). Nevertheless, the SP group performed only slightly better than the IP group in comprehension and oral word recognition. The IP group even scored a little higher than the SP group in connected oral reading."[105]

Response: Garrison and Heard reported scores *separately* for bright and dull students. We did the same.[106] In effect, scores were blocked by levels of intelligence. Total classes were compared in the Garrison and Heard study, but the relevant comparisons are between bright IP and bright SP groups and between dull IP and dull SP groups. That is what is reported in *The Great Debate*, as it should have been. Consequently, the ratios with the regard to intelligence ar irrelevant.

Tate et at. (1940)[107]

Carbo: "In the middle of his study, Harry Tate and his colleagues discovered that the phonics group had a mean mental age that was six months higher than that of the nonphonics group. The difference was nearly identical to the gain differences (both of which favored the SP group) between the two groups on word meaning (6.5 months) and paragraph meaning (7.5 months). The scores of IP and SP groups on word recognition and connected oral reading (which were nearly equal) were not reported by Chall."[108]

Response: The mean mental ages of the IP and SP groups were equivalent when the students were tested in March of first grade. It was not until the posttesting of the same students at the end of grade 2 that differences in mental age between groups appeared. Whatever the reason for the measurement differences noted for the end of grade 2, the differences in intelligence were smaller than is suggested in the *Kappan* article. The mean I.Q. for the IP group was 108.5 and for the SP group, 115 – a difference of 6.5 points. To quote Tate and his colleagues:

> It appears to be established beyond the realm of conjecture that incidental phonics is a much better method than non-phonics in the teaching of primary reading so far as the comprehension of words and paragraphs is concerned. It is hardly reasonable to suppose that superior performances [in reading] as great as 6.5 months and 7.5 months could be attributed to a difference in median intelligence score of 6.5 points.[109]

It is true that Tate and his colleagues reported other findings that were not listed in *The Great Debate*. Since the mean scores on three subtest of the Gates test were near the ceiling for that test, they may have underestimated effects. Therefore, it was proper not to include them. According to Tate and his colleagues, the lack of differences was due to the fact that there was no room left in the test to show differences, not because there was attempt to hide something, as the *Kappan* article implies.

Unequal Instructional Time

Mosher and Newhall (1930)[110]

Carbo:: "Both the IP and SP groups . . . were given identical basal reader instruction. However, the SP group received 15 minutes per day of phonics instruction in addition. Although the students in the SP group received more instructional time, they did not perform significantly better in reading than those in the IP group."[111]

Response: First, Raymond Mosher and Sidney Newhall compared look-say (not intrinsic phonics as the *Kappan* article erroneously stated) with systematic phonics. And, yes, the authors noted that the SP group had 15 minutes more time for phonics. How much effect 15 minutes per day would have remains an open question. Another open question is what the control (look-say) children were doing during that 15 minutes. Also, although they did not perform *significantly* better, the SP group did perform better.[112]

Russell (1943)[113]

Carbo: "In David Russell's study, the SP group had more instructional time in reading and achieved higher reading scores at the end of first grade than did the IP group."[114]

Response: Russell's description does not suggest the large differences in time allotment claimed in the *Kappan* article. Instead, he focused on the instructional differences. A meta-analysis that we performed found very large effect sizes for the Russell study. On some measures, the "much phonics" children outperformed the "little phonics" children by as much as two or three standard units, a very large difference. This could not be accounted for by the small differences in time mentioned in the *Kappan* article.

Summary

Overall, it appears as if the concerns expressed in the *Kappan* article are unfounded. Thus the "verdict" for each of the individual classroom comparisons in *The Great Debate* remain essentially the same.

I would also refer readers to the full discussion of these studies in *The Great Debate*, including the tables on pages 106 to 111. These tables were designed to show the developmental effects of phonics by grade level and reading components. Unfortunately, the "lumping" of the studies in the *Kappan* article[115] by the names of the investigators makes it virtually impossible to see such trends.

Responses to Critical Comments of Specific Studies (1983 Edition)

I present below the criticisms of the studies analyzed in the 1983 updating of *The Great Debate* by categories used in the *Kappan* article (e.g., "nine laboratory experiments"). My responses follow each criticism.

Nine Laboratory Experiments[116]

Carbo: Because the questions are so extensive and varied, rather than quote the *Kappan* article at great length, I refer interested readers to Carbo's full statement on pages 235 and 236 of the November 1988 *Kappan*.

Response: *The Great Debate* stressed that, when evidence from different kinds of studies, with different strengths and weaknesses, all points in the same direction, it should be heeded. For though it is easy to tear individual studies apart, ignoring the whole picture is not wise. Thus laboratory studies, which may not be the best evidence for the effects of phonics, were included but reported separately.

The observation in the *Kappan* article that no study compared phonics to a method stressing meaning misses the point. The point is not whether the words were meaningful or not. What is important is whether certain kinds of systematic attention to the letters and sounds in words leads to better achievement than does no attention, less attention, or a different kind of attention. For these questions, the laboratory studies are certainly relevant.[117]

Eight Classroom Experiments[118]

Carbo: "Chall reported the results of only eight classroom experiments for the period from 1964 to 1981. (Three of the eight were part of the 27 USOE studies.) The eight experiments were conducted over one- to three- year periods and compared the effects of more versus less phonics instruction with primary children. Generally, the finding tabulated by Chall showed some advantage for the more-phonics groups in word recognition and decoding; they showed mixed results in vocabulary and comprehension."[119]

Response: These studies did not, as the *Kappan* article stated, compare the effects of different amounts of phonics. The comparisons were of different kinds of phonics – direct-synthetic as opposed to indirect-analytic.[120] Thus smaller differences could be expected since the earlier comparisons were based on greater differences – no phonics (look-say) to medium amounts (intrinsic phonics) to larger amounts of phonics (systematic phonics).

Readers should also note that the facts in the *Kappan* article are hidden in footnotes, while the negative connotations are stated in the text, without qualifications. The author's comments are also misleading because they confound the findings by treating the results for the different grades as one lump sum, even though a developmental interpretation was central to the analysis.

Carbo: Footnote 77 on page 240 notes that "Chall reported significantly higher scores in vocabulary for the more-phonics groups at the end of third grade in the 1969 Haye and Wuerst study. Hayes and Wuerst did not mention vocabulary results in their 1969 paper, which discussed third-grade findings. In addition, significantly higher oral-reading scores that favored more phonics for th high-I.Q. group in the 1967 Hayes and Wuerst study were tabulated in the update of *the Great Debate*, but significantly higher comprehension and vocabulary scores that favored less phonics for the high group at the end of second grade in the experiment by Putnam and Youtz were unreported by Chall."

Response: True, the vocabulary results were not reported in the paper by Robert Hayes and Richard Wuerst that appears in the bibliography of the update. These results were cited, however, in a paper by the same authors that was published in *Reading Teacher*.[121]

If the findings of Lillian Putnam and Adella Youtz had been included, they would have given additional support to one of the hypotheses in *The Great Debate*: That differences in amount and in type of phonics have less effect on better readers than on poorer readers.

Indeed, Putnam and Youtz made a similar point with regard to the different effects of the two kinds of phonics for urban disadvantaged students. They found that the indirect-analytic phonics group scored higher than the direct-synthetic group in grade 1. However, by grade 2 the indirect-analytic phonics group fell behind the national norms, while the direct-synthetic phonics group gained in grade 2. The authors noted that "this trend may indicate the ultimate advantage for these children of the more thorough approach and mastery of basic skills afforded by the direct phonics program."[122] In this connection, see similar findings by Peter Freebody and Brian Byrne.123

Five Experiments with Exceptional Students[124]

Carbo: "Most of the experiments in this group compared teacher-directed phonics instruction to a linguistic method taught with programmed workbooks (the later taught word families, such as *bet*, *set*, and *met*). Neither method emphasized meaning. Two of the five experiments had design flaws, and one of these two presented data different from those reported by Chall." [125]

Response: The statement about meaning seems to have no relevance here. As I noted in Part 4 above, all methods pay attention to meaning, more or less. The question studied here (as in the laboratory and classroom studies) was, Which kind of phonics is better? These studies provided some useful evidence for answering that question.

Carbo: Two of the studies had designed flaws,[126] those of Joanna Williams and of Catherine Biggins and Sayre Uhler. Williams, the *Kappan* article claimed, equated children in the experimental and control groups "only on their knowledge of phonics."[127] The Biggins and Uhler study had design flaws because "the two groups under study were not equated at the experiment's inception except in that they had 'similar backgrounds.' "[128]

Response: My response to these concerns is that, since the Williams posttests also assessed only knowledge of phonics, equating the experimental and control groups "only on their knowledge of phonics" was appropriate.

As for the Biggins and Uhler study, the groups were not only equated by background, but also reading pretest scores were presented, showing that, at the pretest, the control group had an advantage of .43 of a grade level over the experimental groups. On the posttest, the experimental group had an advantage of .52 of a grade level. In a year's program, the experimental group made about a year's additional growth above that made by the control group. This is a good-size gain, given the exceptional population. It is even greater given the different times for pretesting (see below).

There were some differences in the materials used by the groups, but, except in the phonics programs, there was a very large overlap, as well. All the materials, with the exception of the decoding program, were also used by the control group. The control group did use some additional materials. If the experimental group had had, in addition to the phonics program, a heavy supplement, that would have been a confounding factor. However, since the reverse is true, there does not seem to be a problem. Finally, because of the focus on *beginning* reading, only the second-grade results were examined. The use of different tests is not a serious problem. all students received the same measure as a pretest and posttest. An examination of the raw pretest-posttest gains tells the story:

Group	Pretest	Posttest	Gain
Vocab.			
Exp.	1.29	2.34	1.05
Cont.	1.73	1.82	0.09
Comp.			
Exp.	1.73	2.41	0.68
Cont.	1.71	1.91	0.20

The experimental group was clearly ahead.

One "design flaw" that the *Kappan* article did not point out is that the control groups received their pretesting three months earlier than the experimental groups. Thus the gains reported above represent between seven and eight months for the control groups but only four months for the experimental groups. Once again, this suggests that the direct-synthetic phonics approach had a powerful effect.

The *Kappan* article says that Biggins and Uhler did not find significant differences in vocabulary, while I said that they did. With the use of different tests, testing for statistical significance is probably not the best measure of the effects of the program. But the F ratio of 3.21 closely approached statistical significance. Even without statistical testing, however, the shape of these results should be clear.

In sum, the charges in the *Kappan* article about the shortcomings of the studies reviewed in the update have as little substance as the charges lodged against the studies in the first edition of *The Great Debate*. Thus the overall findings still hold. ◆

References

1. Marie Carbo, "Debunking the Great Phonics Myth," *Phi Delta Kappan*, November 1988, pp. 226-40.
2. Jeanne S. Chall, *Learning to Read: The Great Debate* (New York: McGraw-Hill, 1967); and idem, *Learning to Read: The Great Debate*, updated. (New York: McGraw-Hill, 1983), The original edition of *Learning to Read: The Great Debate*, and inquiry into the science, art, and ideology of old and new methods of teaching children to read, covering the years from 1910 to 1965, was published in 1967. In 1983 an updated edition was published. It contained a 52-page update that reviewed the relevant research published between 1966 and 1982 on the major issues of *The Great Debate*. The remainder of the updated edition consisted of the complete first edition of 1967. Thus the references to the first edition can also be found in the updated edition. When I refer to the new material in updated edition, I specify the 1983 edition.
3. Carbo, p. 226
4. Richard C. Anderson et al., *Becoming a Nation of Readers: The Report of the Commission on Reading* (Urbana: Center for the Study of Reading, University of Illinois, 1985).
5. Carbo, p. 226
6. *The Reading Report Card: Progress Toward Excellence in Our Schools; Trends in Reading our Four National Assessments, 1971-1984* (Princeton, N.J.: National Assessment of Educational Progress and Educational Testing Service, 1985). Moreover, the higher scores of the 9-year-olds were maintained when the cohort was tested again at ages 13 and 17.
7. Helen M. Popp, "Current Practices in the Teaching of Beginning Reading," in John B. Carroll and Jeanne S. Chall, eds., *Toward a Literate Society* (New York: McGraw-Hill, 1975), pp. 101-46.
8. Arthur N. Applebee, Judith A. Langer, and Ina V.S. Mullis, *Learning to Be Literate in America* (Princeton, N.J.: National Assessment of Educational Progress and Educational Testing Service, 1987), p. 35.
9. Dale M. Willows, Diane Borwick, and Maureen Hayvren, "The Content of School Readers," in G. E. MacKinnon and T. G. Waller, eds., *Reading Research: Advances in Theory and Practice*, vol. 2 (New York: Academic Press, 1981).
10. In addition, more resources were available for children at risk – for disadvantaged students in Head Start and for school-age children in Title 1 (later Chapter 1) programs. Children with learning disabilities, at all social levels, received remedial help under P.L. 94-142. "Sesame Street" and "The Electric Company" provided stimulation for learning to read in the home. (See Chall, *Learning to Read: The Great Debate*, 1983, p. 45.)
11. *The Reading Report Card* . . . ., pp. 9, 11.
12. Edward Haertel et al., *Report of the NAEP Technical Review Panel on the 1986 Reading Anomaly, the Accuracy of NAEP trends, and Issues Raised by State-Level NAEP Comparisons* (Washington, D.C.: National Center of Education Statistics and U.S. Department of Education, forthcoming).
13. Jeanne S. Chall, "Could the Decline Be Real? Recent Trends in Reading Instruction and Support in the U.S.," in Haertel et al., op. cit.
14. The most recent NAEP report on reading notes "a greater emphasis on comprehension strategies" in the 1980s. See Arthur N. Applebee, Judith A. Langer, and Ina V.S. Mullis, *Who Reads Best: Factors Related to Reading Achievement in Grade 3, 7, and 11* (Princeton, N.J.: National Assessment of Educational Progress and Educational Testing Service, 1988).
15. Linda A. Meyer, C. Nicholas Hastings, and Robert L. Linn, *Assessing Early Reading with New Decoding and Comprehension Measures* (Urbana: Center for the Study of Reading, University of Illinois, 1988).
16. Nancy Neill, "Analysis of Basal Readers for District Textbook Adoptions," unpublished paper, Graduate School of Education, Harvard University, Cambridge, Mass., 1987.
17. It should be noted that this question is still being researched. See, for example, Haertel et al., op. cit.
18. Jeanne S. Chall, *Stages of Reading Development* (New York: McGraw-Hill, 1983).
19. Jeanne S. Chall, "Reading and Early Childhood Education: The Critical Issues," *Principal*, May 1987, pp. 6-9.
20. The study was conducted from 1962 through 1965 under a grant from the Carnegie Corporation of New York, by a small staff and with the advice of leading scholars in reading and related fields, including John B. Carroll, Arthur I. Gates, Albert J. Harris, Edgar Dale, Helen M. Robinson, C.C. Fries, Nancy Larrick, Allan Barton, David Wilder, M.D. Vernon, Joyce Morris, and John Downing.
21. Chall, 1967, p. 1.
22. Ibid., pp. 100-101.
23. Ibid., p. 176.
24. Ibid., p. 307.
25. C.W. Valentine, "Experiments on the Methods of Teaching Reading," *Journal of Experimental Pedagogy*, vol. 2, 1913, pp. 99-112.
26. Chall, 1967, p. 106.
27. Russell G. Stauffer, "The Verdict: Speculative Controversy," *Reading Teacher*, vol. 19, 1966, pp. 563-64, 575.
28. Guy L. Bond and Robert Dykstra, "The Cooperative Research Program in First-Grade Reading," *Reading Research Quarterly*, vol. 2, 1967, pp. 5-142.
29. Robert Dykstra, "The Effectiveness of Code- and Meaning-Emphasis Beginning Reading Programs," *Reading Teacher*, vol. 22, 1968, pp. 17-23.
30. Robert Dykstra, "Phonics and Beginning Reading Instruction," cited in Charles C. Walcutt, Joan Lamport, and Glenn McCracken, *Teaching Reading: A Phonic/Linguistic Approach to Developmental Reading* (New York: Macmillan, 1974), p. 397.
31. Reginald Corder, *An Information Base of Reading: A Critical Review of the Information Base for Current Assumptions Regarding the Status of Instruction and Reading Achievement in the United States* (Berkeley, Calif.: Educational Testing Service, Berkeley Office, U.S. Office of Education Project 0-9031, ERIC ED 054 922, 1971).
32. Paul B. Diederich, *Research 1960-1970 on Method and Materials in Reading* (Princeton, N.J.: Educational Testing Service, ERIC TM Report 22, ERIC ED 072 115, 1973).

33. Ibid., p. 7.
34. Jane Stallings, *Implementation and Child Effects of teaching Practices in Follow Through Classrooms* (Chicago: University of Chicago Press, Monographs of the Society for Research in Child Development, 1975).
35. Carbo, p. 237.
36. Chall, *Learning to Read: The Great Debate*, 1983, p. 12.
37. See, for example, H. Talmadge and Herbert Walberg, "Naturalistic Decision-Oriented Evaluation of a District Reading Program," *Journal of Reading Behavior*, vol. 10, 1978, pp. 185-95. This study found an advantage for direct phonics instruction in grades 1 through 6.
38. Bond and Dykstra, op. cit.
39. Jeannette Jansky and Katrina de Hirsch, *Preventing Reading Failure* (New York: Harper & Row, 1972), pp. 56-57.
40. See, for example, Jonathan Baron and Ian Thurston, "an Analysis of the Word Superiority Effect," *Cognitive Psychology*, vol. 4, 1976, pp. 177-80; Richard L. Venezky and Dominic W. Massaro, "The Rose of Orthographic Regularity in Word Recognition," in Lauren B. Resnick and Phyllis A. Weaver, eds., *Theory and Practice of Early Reading*, vol. 1 (Hillsdale, N.J.: Erlbaum, 1979), pp. 85-107; Donald Shankweiler and Isabelle Y. Liberman, "Misreading: A Search for Causes," in James Kavanaugh and Ignatius Mattingly, eds., *Language by Ear and by Eye* (Cambridge, Mass.: MIT Press, 1972); Charles Perfetti and Thomas Hogaboam, "The Relationship Between Single Word Decoding and Comprehension Skill," *Journal of Educational Psychology*, vol. 67, 1975, pp. 461-69; Rebecca C. Barr, "The Effect of Instruction on Pupil Reading Strategies, " *Reading Research Quarterly*, vol. 10, 1974, pp. 555-82; Charles Perfetti and Alan M. Lesgold, "Coding and Comprehension and Implications for Reading Instruction," in Resnick and Weaver, pp. 57-84; and David Laberge and S. Jay Samuels, "Toward a Theory of Automatic Information Processing in Reading," in Harry Singer and Robert B. Ruddell, eds., *Theoretical Models and Processes in Reading*, 2nd ed. (Newark, Del.: International Reading Association, 1976).
41. Robert C. Calfee and Priscilla A. Drum, "Learning to Read: Theory, Research, and Practice," *Curriculum Inquiry*, vol. 8, 1978, pp. 183-249.
42. Charles Perfetti, *Reading Ability* (New York: Oxford University Press, 1985).
43. See Laberge and Samuels, op. cit.; Keith E. Stanovich, "Mathew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy," *Reading Research Quarterly*, Fall 1986, pp. 360-407; and Peter Freebody and Brian Byrne, "Word-Reading Strategies in Elementary School Children; Relations to Comprehension, Reading Time, and Phonemic Awareness,"*Reading Research Quarterly*, Fall1988, pp. 441-53.
44. Perfetti, p. 231.
45. Stanovich, p. 364.
46. Marcel A. Just and Patricia Carpenter, *The Psychology of Reading and Language Comprehension* (Newton, Mass.: Allyn and Bacon, 1987), p. 51.
47. Ibid., p. 328.
48. Linnea C. Ehri, "Learning to Read and Spell Words," *Journal of Reading Behavior*, vol. 19, 1987, pp. 5-30.
49. Lynette Bradley and Peter Bryant, "Categorizing Sounds and Learning to Read – A Causal Connection," *Nature*, vol. 301, 1983, pp. 419-21; and idem, *Rhyme and Reason in Reading and Spelling* (Ann Arbor: University of Michigan Press, 1985).
50. Keith E. Stanovich, "Introduction," *Merrill-Palmer Quarterly*, vol. 33, 1987, p. ii.
51. Tom Nicholson, Christine Lillas, and M. Anne Rzoska, "Have We Been Misled by Miscues?," *Reading Teacher*, vol. 42, 1988, p. 9.
52. Freebody and Byrne, op. cit.
53. Ibid., p. 441.
54. Ingvar Lundberg, Ake Olofsson, and Stig Wall, "Reading and Spelling skills in the First School Years, Predicted form Phonemic Awareness Skills in Kindergarten," *Scandinavian Journal of Psychology*, vol. 21, 1980, pp. 159-73, cited in Isabelle Y. Liberman, "Language and Literacy: The Obligation of the Schools of Education," in Rosemary Bowler, ed., *Intimacy with Language: A Forgotten Basic in Teacher Education* (Baltimore: Orton Dyslexia Society, 1987), p. 3.
55. Isabelle Y. Liberman and Donald Shankweiler, "Speech, the Alphabet, and Learning to Read," in Lauren B. Resnick and Phyllis A. Weaver, eds., *Theory and Practice of Early Reading*, vol. 2 (Hillsdale, N.J.: Erlbaum, 1979); Joan Helfgott, "Phonemic Segmentation and Blending Skills of Kindergarten Children: Implications of Beginning Reading Acquisition," *Contemporary Educational Psychology*, vol. 1, 1976, pp. 157-69; Barbara Fox and Donald K. Routh, "Phonemic Analysis and Synthesis as Word Attack Skills," *Journal of Educational Psychology*, vol. 68, 1976, pp. 70-74; and Rebecca A. Treiman and Jonathan Baron, "Segmental Analysis Ability: Development and Relation to Reading Ability," in T.G. Waller and G.E. MacKinnon, eds., *Reading Research: Advances in Theory and Practice*, vol. 3 (New York: Academic Press, 1982), pp. 159-98.
56. Joanna P. Williams, "The ABDs of Reading: A Program for the Learning Disabled," in Resnick and Weaver, vol. 2, pp. 179-195; Michael A. Wallach and Lise Wallach, "Helping Disadvantaged Children Learn to Read by Teaching Them Phoneme Identification Skills," in Resnick and Weaver, vol. 2, pp. 197-215; William E. Tunmer and Andrew R. Nesdale, "Phonemic Segmentation Skill and Beginning Reading," *Journal of Educational Psychology*, vol. 77, 1985, pp. 417-27; and Bradley and Bryant, op. cit.
57. Mary Ann Evans and Thomas M. Carr, "Cognitive Abilities, Conditions of Learning, and the Early Development of Reading," *Reading Research Quarterly*, vol. 20, 1985, pp. 327-50.
58. Steven A. Stahl and Patricia D. Miller, "Whole Language and Language Experience Approaches for Beginning Reading: A Quantitative Research Synthesis," *Review of Educational Research Quarterly*, in press.
59. See Chall, *Stages of Reading Development.*
60. Liberman, p. 1.
61. Ibid., p. 5.
62. Just and Carpenter, op. cit.
63. Williams, op. cit.; and Wallach and Wallach, op. cit.
64. Anderson et al., op. cit.
65. Ibid., p. 37.
66. Miriam Balmuth, *The Roots of Phonics* (New York: McGraw-Hill, 1981).
67. Dina Feitelson, *Facts and Fads in Beginning Reading: A Cross-Language Perspective* (Norwood, N.J.: Ablex, 1988).
68. Marilyn J. Adams, "Phonics and Beginning Reading Instruction," draft of a report prepared for the Reading Research and Education Center, University of Illinois, Urbana, 1988.
69. Ibid., p. 1.
70. Ibid.
71. Ibid., p. 20.
72. Ibid.
73. Ibid., p. 21.
74. Jeanne S. Chall, *Reading 1967-1977: A Decade of Change and Promise* (Bloomington, Ind.: Phi Delta Kapa Educational Foundation, Fastback 97, 1977).
75. Joanna P. Williams, "Reading Instruction Today," *American Psychologist*, vol. 34, 1979, pp. 310-15.
76. Stanovich, "Introduction," p. iv.
77. Stanovich, "Matthew Effects. . . . "
78. Carbo, p. 236.

79. Chall, 1967, p. 138.
80. Chall, *Stages of Reading Development.*
81. The average first-grader can understand and use in speech about 6,000 words. It will take at least three years before a first-grader can recognize these words in print. See Chall, *Stages of Reading Development*, pp. 62-63.
82. Nila B. Smith, *American Reading Instruction* (Morristown, N.J.: Silver Burdett, 1934).
83. Charles Read, "Preschool Children's Knowledge of English Orthography," *Harvard Educational Review*, vol. 41, 1971, pp. 1-34.
84. Glenda Bissex, *GNYS AT WRK* (Cambridge, Mass.: Harvard University Press, 1980).
85. For what is perhaps the most popular of the whole-language approaches, see Kenneth Goodman, *What's Whole in Whole Language?* (Portsmouth, N.H.: Heinemann, 1987). For a recent overview, see *ASCD Update*, January 1989, published by the Association for Supervision and Curriculum Development, Alexandria, Va. See also Carbo, fn. 88, p. 240. For sorting out the different methods and approaches, see Appendix A, Chall, "Classification of Twenty-Two Beginning Reading Programs," *Learning to Read: The Great Debate*, 1967 or 1983, pp. 336-55. This will be helpful in understanding the different reading approaches and why they are best viewed not as completely opposite to each other, but as containing certain similar and certain different feature.
86. Carbo, pp. 236-37.
87. Mary Wohleber, "A Study of the Effects of a Systematic Program of Phonetic Training on Primary Reading" (Doctoral dissertation, University of Pittsburgh, 1953).
88. Carbo, p. 229.
89. Chall, 1967, p. 104. My interpretation is further supported by Louise Gurren and Ann Hughes, "Intensive Phonics vs. Gradual Phonics in Beginning Reading: A Review, " *Journal of Educational Research*, vol. 58, 1965, pp. 339-46. They reviewed the first-grade results as follows: "The majority of the relevant subtests, two out of three, showed significant differences favoring the 65 first-graders who had used a modified linguistic program. There were no significant differences favoring the 65 'conventional' first-graders" (p. 346).
90. Mary Edward, "A Modified Linguistic Versus a Composite Basal Reading Program," *Reading Teacher*, vol. 17, 1964, pp. 511-15.
91. Carbo, p. 229.
92. Edward, p. 512.
93. Elmer Morgan, Jr., and Morton Light, "A Statistical Evaluation of Two Programs of Reading Instruction," *Journal of Educational Research*, vol. 57, 1963, pp. 99-101.
94. Carbo, p. 231.
95. Donald C. Agnew, "The Effect of Varied Amounts of Phonetic Training on Primary Reading" (Doctoral dissertation, Duke University, 1935); and idem, *The Effect of Varied Amounts of Phonetic Training on Primary Reading* (Durham, N.C.: Duke University Research Studies in Education, No. 5, Duke University Press, 1939).
96. Carbo, p. 232.
97. See Agnew, 1939, p. 37.
98. Ibid., p. 41.
99. Ibid., p. 47.
100. For a fuller explication, see Chall, 1967, pp. 94 and 302. See also Edward, op. cit.; Freebody and Byrne, op. cit.; and Chall, *Stages of Reading Development.*
101. Arthur I. Gate, "Studies of Phonetic Training in Beginning Reading," *Journal of Educational Psychology*, vol. 18, 1927, pp. 217-66.
102. Carbo, p. 232.
103. Gates, p. 219.
104. S.C. Garrison and M. T. Heard, " An Experimental Study of the Value of Phonetics," *Peabody Journal of Education*, vol. 9, 1931, pp. 9-14.
105. Carbo, p. 233.
106. See Chall, 1967, Table 4-2A, p. 109.
107. Harry L. Tate, Theresa M. Herbert, and Josephine K. Zeman, "Nonphonic Primary Reading," *Elementary School Journal*, vol. 40, 1940, pp. 529-37.
108. Carbo, p. 233.
109. Tate, Herbert, and Zeman, p. 534.
110. Raymond M. Mosher and Sidney M. Newhall, "Phonic Versus Look-and-Say Training in Beginning reading," *Journal of Educational Psychology*, vol. 21, 1930, pp. 500-506.
111. Carbo, p. 233.
112. See Chall, 1967 or 1983, Table 4-1B, p. 107.
113. David H. Russell, "A Diagnostic Study of Spelling Readiness," *Journal of Educational Research*, vol 37, 1943, pp. 276-83.
114. For a full account, see Carbo, p. 233.
115. Carbo, pp. 230-31.
116. The nine laboratory studies were: Douglas W. Carnine, " Phonics vs. Look-Say: Transfer to New Words," *Reading Teacher*, vol 30, 1977, pp. 636-40; Barbara J. Fox and Donald K. Routh, "Phonemic Analysis and Synthesis as Word-Attack Skills," *Journal of Educational Psychology*, vol. 68, 1976, pp. 70-74; Maryann Haddock, "The Effects of an Auditory and Auditory-Visual Method of Blending Instruction on the Ability of Prereaders to Decode Synthetic Words," *Journal of Educational Psychology*, vol. 68, 1976, pp. 825-31; idem, "Teaching Blending in Beginning Reading Instruction Is Important," *Reading Teacher*, vol. 31, 1978, pp. 654-58; W.S. Jeffrey and S.J. Samuels, "Effect of Method of Reading Training on Initial Learning and Transfer," *Journal of Verbal Learning and Verbal Behavior*, vol. 6, 1967,, pp. 354, 358; Joseph R. Jenkins, R. Barker Bausell, and Linda M. Jenkins, "Comparison of Letter Name and Letter Sound Training as Transfer Variables," *American Educational Research Journal*, vol. 9, 1972, pp. 75-86; Douglas Muller, "Phonic Blending and Transfer of Letter Training to Word Reading in Children," *Journal of Reading Behavior*, vol. 5, 1972-73, pp. 212-17; Thomas R. Vandever and Donald D. Neville, "Transfer as a Result of Synthetic and Analytic Training Instruction," *American Journal of Mental Deficiency*, vol. 80, 1976, pp. 498-503; and Thomas D. Yawkey, "Reading Training and Rural Disadvantaged Five Year Old Children," *Reading World*, vol. 13, 1973, pp. 128-40.
117. See Chall, *Learning to Read: The Great Debate*, 1983, Table 1-1, "Laboratory Type Experiments," p. 14.
118. The eight classroom experiments were: Ernest M. Grant, "A Study of the Comparison of Two Reading Programs (Ginn 360 and DISTAR) upon Primary Urban Inner City Students" (Doctoral dissertation, University of Wisconsin, 1973); Robert B. Hayes and Richard C. Wuerst, "ITA and Three Other Approaches to Reading in the First Grade - Extended into Second Grade," *Reading Teacher*, vol. 20, 1967, pp. 694-98; idem, "Four Instructional approaches to Beginning Reading - Three Years Later, " paper presented at the annual convention of the International Reading Association, Boston, 1969 (ERIC ED 020 098); LaVor Lym, *Basal Reading Program: Keys to Reading* (Dallas: Department of Research and Evaluation, Dallas Independent School District, Research Report No. 73-144, 1973); Pittsburgh Learning Research and Development Center, "Report of the First Year Implementation of the New Primary Grades Reading System (NRS) in Kelly School. (1974-1975 School Year)," mimeo, n.d.; Marian Potts and Carl Savino, "The Relative Achievement of First Graders Under Three Different Reading Programs," *Journal of Educational Research*, vol. 61, 1968, pp. 447-50; Lillian R. Putnam and Adella C. Youtz, "Is a Structured Reading Program Effective for Disadvantaged Children?" *Reading World*, vol. 12, 1972, pp. 123-35; and Nita M. Wyatt, *Reading Achievements of First Grade Boys Versus First Grade Girls Using Two Approaches: a Linguistic Approach and Basal Reader Approach with Boys and Girls Grouped Separately* (Lawrence: University of Kansas, Cooperative Research Project No. 2735, 1965).
119. Carbo, p. 236.
120. See Chall, *Learning to Read: The Great Debate*, 1983, Table 1-2, pp. 18-19.

Book Review—
Action Research in the Secondary School

by R. Paul Gregory, Educational Psychologist,
City of Birmingham (U.K.) Education Department
Published by Routledge, London and New York, 1988.
Price $45.00 US, $58.50 Can. Pages 185.

Action Research describes the personal odyssey of School Psychologist Paul Gregory, over a fifteen-year career, by bringing together 10 of his previously published papers. The book is aimed at showing that a psychologist working in schools could remain a scientist and in the process contribute to better decision making in schools. The book should make good reading for school psychologists in training, for those who train school psychologists, and school administrators looking for better use of their psychologists in decision making.

A first study sought an easy way to identify students "at risk" for failure. Gregory applied a scale proposed by Herbert (1977) which included: (a) quality of clothing (b) attendance, (c) parental occupation, (d) family size, and (e) amount of parent contact with school. In a high school for girls, 15 of the 123 first-year students (12.2%) were identified as socially disadvantaged by this scale and were found more likely to be poor readers and poorly motivated for school. By early identification of these students, it was hoped that ways could be found to improve their chances of success.

A subsequent use of this "at risk" scale in a more disadvantaged high school (22%) was carried out and the results compared with those from the first study. There were even more poor readers and poorly motivated students. The analysis showed that first year students pulled out for remedial work had the worst attendance records and the poorest parental involvement. Latter studies using SRAs *Corrective Reading Programs* at this school (also covered in the book) showed how Gregory used data to pursue solutions to the problem of improving the school involvement of the parents of "at risk" students. An action research plan was developed to promote increasing parent teacher communications. Several studies were carrried out to explore procedures that worked—and some did! Another action research was aimed at increasing attendance by decreasing truancy. Finally, two chapters report evaluations of the use of SRAs *Corrective Reading Program*, along with related research by others. The improvements in reading outcomes obtained with CRP were quite dramatic.

A concluding section examines what he had learned. Gregory found that school personnel were very willing to participate in action research. He found it best to focus on one school and look for what needed to be done that could contribute to its serving students better. It was also helpful to work with several teachers at a time so they could then provide mutual support to each other. He found his studies didn't get written up unless he planned explicit time in his schedule for such activities. These observations are very consistent with our own observations from doing action research in the schools.

From the point of view of experimental findings that could be widely generalized to other settings, these studies all have obvious limitations, as might be expected. Paul Gregory is as aware of this as anyone and cautiously draws his conclusions. However, as a model of what psychologist can do to help school systems and their students, the book is exciting. It fits in with the work on Precision Teaching, Curriculum Based Assessment, and Direct Instruction in providing models for the better uses of psychologists in schools. ◆

Reviewed by
Wesley C. Becker

"Phonics Myth"—*Continued*

121. Hayes and Wuerst, "ITA and Three Other. . . . "
122. Chall, *Learning to Read: The Great Debate*, 1983, p. 15.
123. Freebody and Byrne, op. cit.
124. The five experiments with exceptional populations were: Catherine M. Biggins and Sayre Uhler, "Is There a 'Workable' Decoding System," *Reading Improvement*, vol. 16, 1979, pp. 47-55; Mary Lee Enfield, " An Alternative Classroom Approach to Meeting Special Learning Needs of Children with Reading Problems" (Doctoral dissertation, University of Minnesota, 1976); Ellis Richardson, Bertrand G. Winsberg, and Irv Bialer, "Assessment of Two Methods of Teaching Phonic Skills to Neuropsychiatrically Impaired Children," *Journal of Learning Disabilities*, vol. 10, 1973, pp. 628-35; David A Sabatino and Nancy Dorfman, "Matching Learner Aptitude to Two Commercial Reading Programs," *Exceptional Children*, vol. 41, 1974, pp. 85-90; and Joanna P. Williams, "Teaching Decoding with an Emphasis on Phonemic Analysis and Phoneme Blending," *Journal of Educational Psychology*, vol. 72, 1980, pp. 1-15.
125. Carbo, p. 236.
126. See Carbo, p. 236 and fns. 79 and 80, p. 240.
127. Carbo, p. 240.
128. Carbo, fn. 79, p. 240.

Decoding Words: The Facilitative Effects of Saying the Sounds in a Word—Without Pausing*

by Paul Weisberg
University of Alabama
Tuscaloosa, Alabama

Introduction

There is an alarming and growing number of children who go through school without being able to decipher new words. This inability to decode words has an immediate deleterious effect on word recognition and, eventually, it precludes reading comprehension (Anderson, Hiebert, Scoot, & Wilkinson, 1985). Given these difficiencies, it is not uncommon for many new readers to withdraw from dealing with text altogether.

The skills for decoding printed words, according to one approach (Carnine & Silbert, 1979), involves the two component processes depicted in Figure 1.

Segmenting skills are involved when words are sounded-out by saying the sounds discretely in a left-to-right order. *Blending* skills are used to recombine the subdivided units to arrive at the words at its normal spoken rate. For example, after initially segmenting *mad*, either as *mm–aa–d* or *mmaad*, the reader during the blending component would be expected to say *mad!*

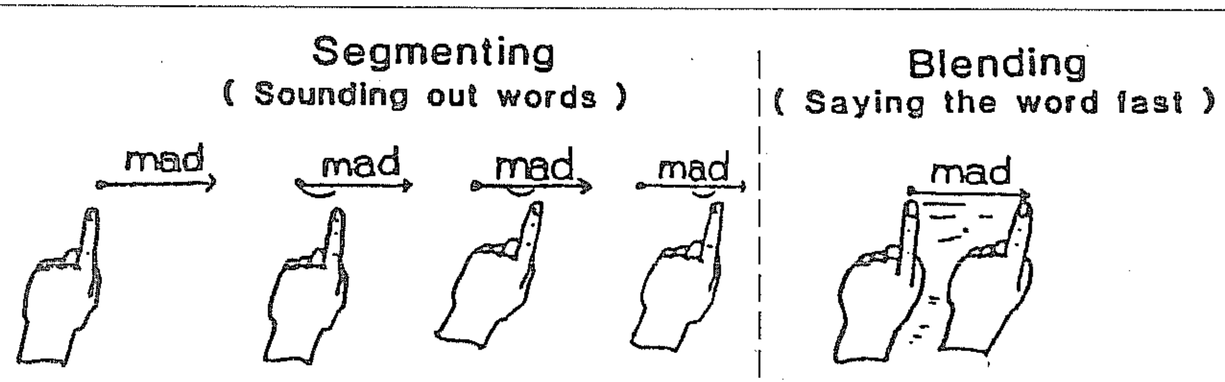
* To be published in *Education and Treatment of Children*. Reprinted with permission of the publisher and the author. Copyright 1990.

In many phonic-based programs, the segmented units are said with silent pauses of unknown duration between them as in "mm (pause) . . . aa (pause) . . . d." The apparent justification for interjecting pauses between sounds is to facilitate the segmenting process: to demonstrate to the naive reader that words are divisible into parts. However, if pausing serves this discriminative function, it could be more than offset by the behavorial problems it creates during blending. Having paused between each sound, the reader could now find it extremely difficult to recombine the interrupted sounds in order to produce the word.

Research on oral blending (Weisberg, Andracchio, & Savard, 1989) found that when an adult narrator paused between successive sounds in one-syllable words for as little as one second, the percent of correctly blended words for nonreading kindergartners was a dismal 16% and only 50% for first graders. When the narrator did not pause between sounds, the levels were 60% and 73%, respectively, for the two groups. Direct training in oral blending also made a difference: kindergartners accustomed to hearing a teacher segment words without pausing were 85% correct in identifying both familiar words and pseudowords.

The effect of pausing between sounds when decoding printed words has not been investigated. Improved accuracy is expected if a strategy is taught that directs new readers to sound-out words without pausing. Saying the sounds continuously without pausing

Figure 1. Segmenting and blending of printed words. During *segmenting*, the individual (teacher or student) begins by touching the ball to the left and then loops from letter-to-letter saying each sound either by not pausing between them (No-Pause Condition) or by pausing 1-second between them(1-Second Pause Condition). During *blending*, the individual touches the ball and slashes under the entire word, saying the word at it's spoken rate—mad!



should make the segmented but blended word functionally more like the actual word (*mmaad* and *mad!*) than when pausing is involved (*mm-aa-d* and *mad!*). There is the increased probability that nonpausing during segmenting will enable the reader to come up with the correct word and, conversely, pausing should invite hesitations, word distortions, and guessing.

Given these possibilities and the goal of behavioral analysts in education to seek more effective technologies for all children, the present experiment led to a long-term analysis of the functional effects of pausing on the decoding process with at-risk preschool children.

Method

Subjects. Nine preschoolers, four to five years old, who attended a preschool for mostly poverty-level children (Weisberg, 1988) participated. Upon entry, non could read, spell, or do simple arithmetic computation.

The reading program. The first 90 lessons of the Reading Mastery I Series (Engelmann & Bruner, 1983), which teaches decoding strategies of the kind shown in Figure 1, were followed.

Training sequence. The chronology of training can be summarized as follows:

A. *Single-group training.* All children attended the same reading class for 30 lessons. Common prerequisite skills for decoding were taught: identifying sounds, left-to-right sequencing, and orally segmenting and blending of words.

In addition, segmenting and blending several two-letter words (e.g., *ma, am, me, se*) without pausing between sounds was taught.

B. *Decoding training with and without pausing.* Two groups were formed. For the No-Pause (0-sec.) Condition ($N = 5$), the teacher did not pause between each sound. In the 1-Second Condition ($N = 4$), the teacher paused one second between each sound.

C. *Introducing the No-Pause Condition for 1-Second trained children.* After a variable number of reading lessons, selected children from the 1-Second Pause classroom were placed in the No-Pause classroom for remediation purposes.

Teaching procedure. A three-step instructional procedure was generally followed: (1) teacher *modeling* of the experimentally appropriate segmenting and blending strategies; (2) teacher *leading,* whereby the teacher and children did segmenting and blending activities together; and (3) group or individualized *testing,* whereby the teacher called on the entire group or individual children and tested them on how well they decoded the target words. For the No-Pause Condition, the teacher modeled and led by not stopping between the sounds whereas in the 1-Second Pause Condition, the teacher modeled and led by pausing one second between successive sounds. During testing, the teacher reinforced or corrected the children in the use of the response typographies appropriate to each pause condition.

Whether in the same or in different classrooms, all children were taught the common sounds for letters as well as skills dealing with rhyming words, reading one- and two-sentence stories, picture-word comprehension tasks, and printing manuscript letters.

Performance Measures. The children were individually evaluated without feedback across 12 weekly scheduled probe sessions on a sample of sounds and words from lessons taught that week as well as from previous lessons. The three performances measured during each probe were the ability:

(1) to correctly say the sounds for recently trained and familiar (reviewed) letters, identified as *sounds* in Figures 2 and 3.

(2) to segment or sound out recently trained and familiar words without stopping between the sounds, identified as *not-stopping* in Figures 2 and 3.

(3) to blend the sounds and come up with the correct word, identified as *words* in Figures 2 and 3.

During the 90 reading lessons, over which the 12 probes sessions were distributed, a total of 21 letter-sounds and 130 words were taught. Early probes typically evaluated performance on 5 sounds and on 6 words whereas the later probes each evaluated 10 sounds and 12 words. The sounds making up the words to be decoded were highly practiced, thus making the words probed in Figures 2 and 3 familiar ones. (Additionally probed were relatively less-familiar words comprising recently taught sounds, but individual performance on these words is not reported here.)

Experimental validity and reliability. Two teachers alternated between the two classes. Periodic evaluation of the teachers by two independent observers showed that on 98% of the intervals recorded, the teacher appropriately modeled and led the children in using the designated pause interval.

Median percent interobserver agreement between two observers was evaluated on the reading performances for at least 10 probe sessions. For sounds it was 98% (83% to 100%), for not-stopping, it was 96% (93% to 100%) and for correct word identification it was 100%. In not- stopping, the end of one sound and the beginning of the next sound had to be said continuously without a distinct pause.

Results

A. *No-Pause* (0-sec.) *Condition* (Figure 2).

High levels of performance are evident for sounds, not-stopping, and (blended) word identification for all preschoolers. Correct word identification did not immediately emerge in the early probes even though

Figure 2. Probe performance of preschoolers trained through out to decode under the No-Pause (0-sec) Condition.

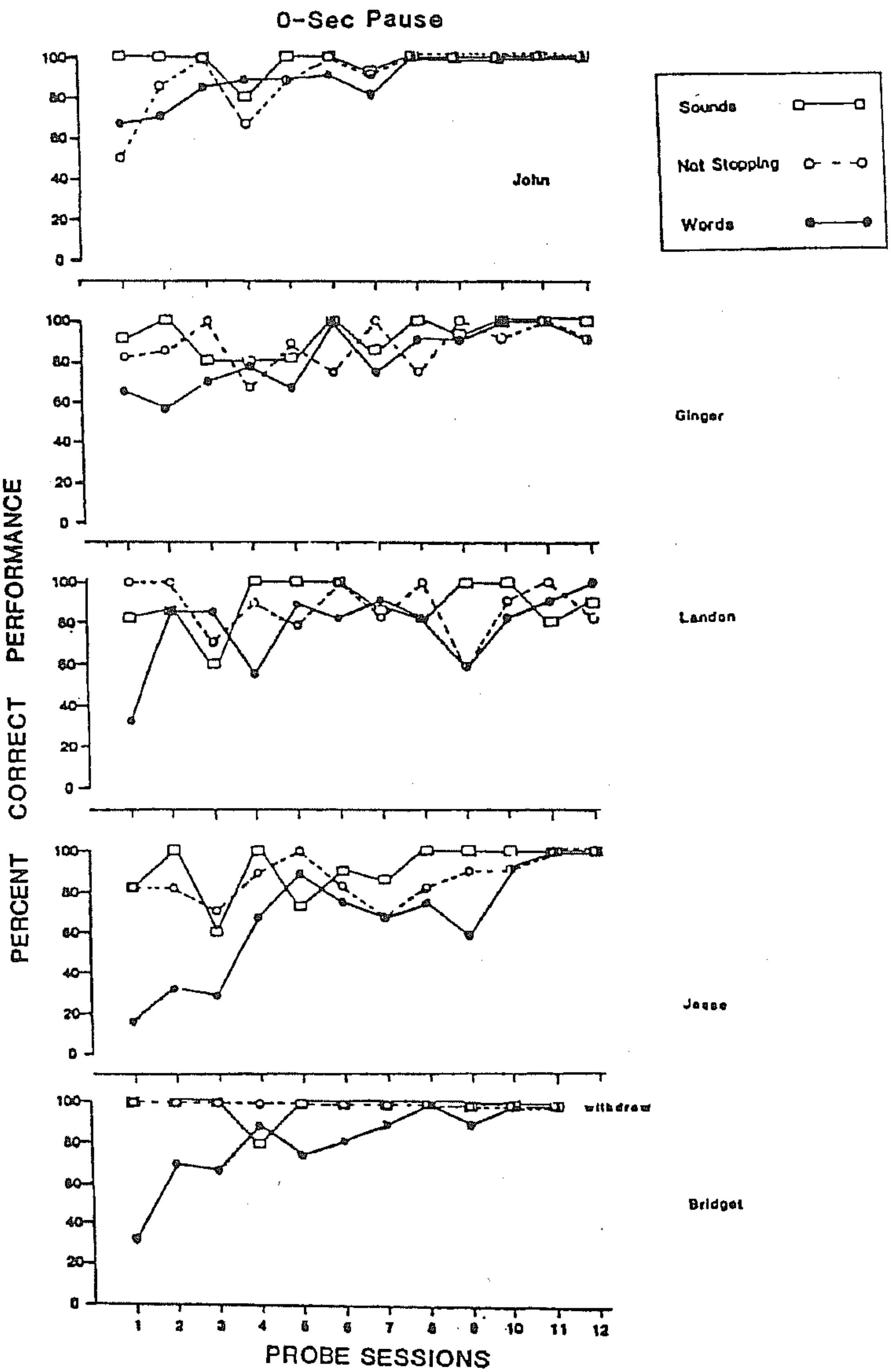
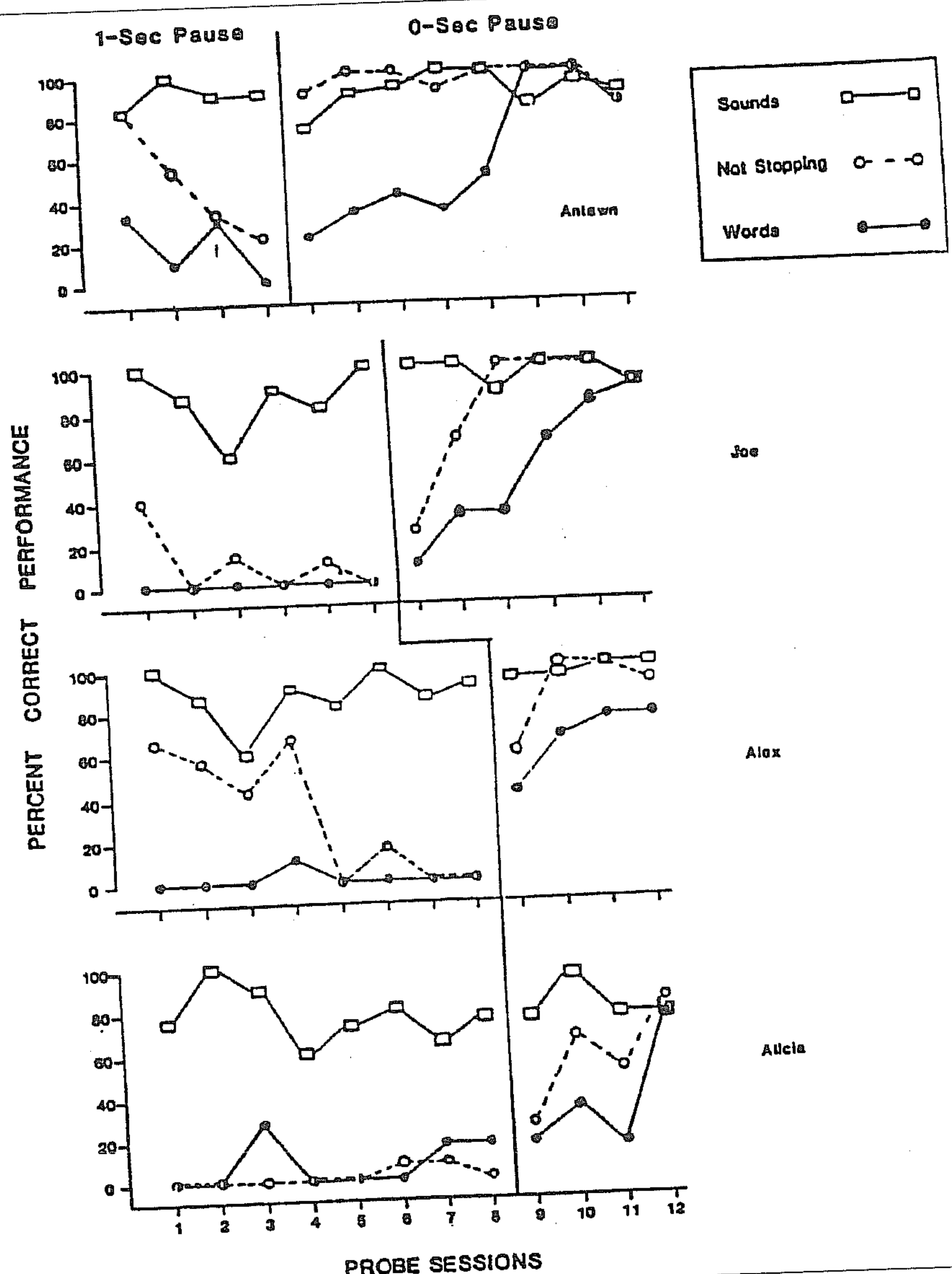


Figure 3. Probe performance of preschoolers trained to decode initially under the 1-Second Pause Condition and then switched to the No-Pause (0-sec) Condition.



not-stopping performance had been elevated. Apparently, a build-up of classroom practice in not-stopping was necessary to ensure that the children would continue to use this strategy to decode words.

B. *One-Sec. Pause to No-Pause Conditions* (Figure 3).

Performance in not-stopping and word identification was either at low levels during the first few probes or it worsened across the probe sessions for the children trained at the 1-sec pause condition. Possibly, the early elevated levels of not-stopping performance by Antawn and Alex was due to their history of being taught to sound out two-letter words without pausing. Sound identification was largely about 80%, save for Alicia.

The No-Pause Condition appeared to have remedial effects, although the benefits were not immediate. Again, the development of not-stopping behavior appeared necessary for appropriate decoding of words to take hold.

Not shown is the fact that children receiving extensive practice with the 1-Second Pause Condition (Alex and Alicia) when switched to the No-Pause Condition required three times as many catch-up reading lessons as those receiving earlier remediation (Antawn & Joe). The former children had to be taught how to decode not only the recently taught words scheduled for the upcoming probe but, also, all of the words previously evaluated. Once retrained on the beginning words, they were *re-probed*, the results of which (not shown in Figure 3) revealed continued decoding improvement in the earlier re-probed words.

The necessity of the many catch-up lessons at the No-Pause Condition and the slow development of decoding suggested by the re-probed data argues strongly for the teaching of decoding skills without stopping between the sounds form the very beginning of training.

Conculsions

Teaching beginning readers to decode words by saying the sounds without pausing between each one is a far more effective procedure than by pausing between the sounds. No-pausing training was associated with more familiar words being correctly decoded. The poor reading abilities of children originally taught by no-pausing between sounds could be over-come if they were remediated by not-pausing training. The extent of remediation can be expected to depend upon the length of prior, ineffective decoding training produced by pausing between sounds. ◆

References

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The Report on the commission on reading*. Washington, D.C.: The National Institute of Education, U.S. Department of Education.

Carnine, D., & Silbert, J. (1979). *Direct instruction reading*. Columbus, OH: Merrill.

Engelmann, S., & Bruner, E. (1983). *Reading Mastery I, DISTAR Reading*. Chicago: Science Research Associates.

Weisberg, P. (1988). Direct Instruction in the preschool. *Education and Treatment of Children, 11*, 349–363.

Weisberg, P., Andracchio, B. J., & Savard, C., F (1989). Oral blending in young children: Effects of pausing between sounds, type of initial sound and word familiarity. *Journal of Educational Research, 82*, 139–145.

Oral Blending in Young Children: Effects of Sound Pauses, Initial Sound, and Word Familiarity*

by Paul Weisberg
Bruno J. Andracchio
Christopher F. Savard
University of Alabama

In an oral blending task, an individual is told the parts of a word (subdivided into a series isolated sounds, sound combination, syllables, or other phonological components) and is expected to produce the word form these constituents. Many researchers acknowledge that practice in oral blending can facilitate the accuracy of decoding printed words (Carnine & Silbert, 1979; Richardson, DiBenedetto, & Bradley, 1977; Rohnback, Bell, & McLaughlin, 1982). This assumption is substantiated by the modest correlation (.40 to .60) found between blending and concurrent measures of word recognition taken from the first to the fourth grades (Chall, Roswell, & Blumenthal, 1963; Williams, 1980).

Oral blending ability also is predictive of later reading (Chall et al., 1963; Goldstein, 1976), and, if reading is taught by phonic approaches, blending may be a more important determinant than IQ (Richardson, DiBenedetto, & Bradley, 1977).

Despite the apparent simplicity involved in recom-

*From *Journal of Educational Research*, Vol. 82, pp. 139-145, 1989. Reprinted with the permission of the Helen Dwight Reid Educational Foundation. Published by Heldref Publications, 4000 Albemarle St., NW, Washington, D.C. 20016. Copyright 1989.

bining or synthesizing an ordered set of spoken sounds to form a word, blending tasks have been particularly troublesome for low-performing readers. For example, low-socioeconomic (SES), inner-city children in Grades 1, 2-3, and 4 could correctly blend only 8%, 25%, and 42%, respectively, of the sounds in CVC words (e.g., r-u-g) (Chall et al. 1963). Williams (1980) assessed blending of CVC words with a mixed-age group of 7- to 12-year-old children receiving remedial reading and found accuracy to be 26% in one study and 40% in another. Ramsey (1972) (cited in Haddock, 1976) noted that 40% of the errors made with unfamiliar words in context by marginal second-grade readers were due to blending difficulties, even though they knew the elements of the word.

Typically, the to-be-blended components, when spoken by an examiner, are said discontinuously and are broke by silent pauses. The learner is expected to arrive at the word by bridging the intervening intervals with relevant phonological information. Pausing between segmented sounds or syllables is a presentation feature inherent not only in developmental and training studies of blending (Chall et al. 1963; Goldstein, 1976; Haddock, 1976; Muller, 1972-1973), but also in standardized assessment procedures, such as the blending test of Katz and Harmon (1982) and Roswell and Chall (1963) and the blending subtests of other instruments (Goldman, Fristoe, & Woodcock, 1976; McCarthy & Kirk, 1961). The dismal blending performance of instructionally naive students and poor readers may be a function of the length of the time interval intervening between the segmented sounds of a word.

In Experiment 1, we tried to ascertain whether pausing between consecutive sounds of a word would be more detrimental to blending accuracy than a never-investigated condition of not pausing between sounds. We studied kindergarten and first-grade children partly to assess the influence of developmental differences on blending. Large differences have been found between those two ages groups in their ability to process and arrange phonetic sequences (Calfee, Lindamood, & Lindamood, 1973) and to add, omit, substitute, or rearrange phonemes (Rosner & Simon, 1971).

We also evaluated the ability to blend words of different semantic value (familiar versus nonsense or nonwords). Evidence exists that blending real words is easier than nonwords (Williams, 1980), but researchers do not know whether word familiarity interacts with the duration of the pause interval. At a 0-sec pause interval, therefore, no performance differences between the two types seem likely, whereas at increasingly longer pause intervals, real words might be easier to blend than nonwords. The decision of whether the initial consonant sound in a CVC word is a continuant (capable of being said continuously, as in *m* or *s*) or a stop (said briefly, as in *p* or *t*) also could affect blendability.

Because saying stop sounds is assumed to invite the addition of an "intrusive vowel" (e.g., *t* as in *tuh*), a greater distortion in blending should result from words beginning with stop compared with continuant consonants (Gleitman &Rozin, 1973). Of the two types, Coleman (1970) found differences in favor of continuants placed at the beginning and end of words, whereas Haddock (1976) reported no differences.

In Experiment 2, we evaluated blending ability in children who were taught to read through an approach that provided extensive training in synthetic phonics. In this replication, we expected that as training in phonics increased so would oral blending for all durations of delay between sounds. Because Experiment 2 participants were preschoolers, many of whom were able to read at advanced levels (Weisberg, in press), the documentation of levels of blending accuracy higher than that normally seen with this age group (Goldstein, 1976; Rosner, 1973-1974) was a possibility.

Experiment 1

Method

Subjects. We selected kindergarten and first-grade children form heterogeneously grouped homerooms that reflected the same composition of SES backgrounds as the entire public school. The school was racially integrated and included 25% of the children from poverty-level homes (eligible for free and reduced-price meals), 40% from blue-collar homes, and 35% from middle-class homes. We conducted the study during March and April when the kindergarteners' mean age was 71.60 months (SD = 5.43) and the first graders' mean age was 84.91 months (SD = 5.18). The kindergarten and first-grade groups contained 15 and 12 females, respectively; and 7 and 12 boys.

The first graders used the Houghton Mifflin basal series, a meaning-emphasis program that teaches whole words with phonic exercises built into the lessons. Skills for oral blending, however, were not taught directly (Beck & McCaslin, 1978). The kindergarteners were not given any formal training, either in oral blending or in reading, and they were considered non-readers.

We obtained parental consent for participation. Two first-grade student did not participate due to lack of consent for one and relocation of the other. One kindergartener had poorly articulated, unclear speech that precluded participation. another was excluded through a random *draw* so that each age group would have 22 children.

Composition of target words. Each of 20 meaningful and 12 nonwords consisted of three distinct phonemes. The meaningful words were: *fish, rag, meat, nose, luck, save, whine, them, not, what, gave, cot, deep, give, hen, joke, kite, pan, tube,* and *pull.* Most of these words are high-frequency terms, ranking high on published word lists for children (Barnard & Degracie, 1976; Dolch, 1936). Because of the desire to sample a broad range of different medial vowels and different beginning and ending consonants, however, some words were sampled that ranked lower in frequency of usage, for example, *luck, whine, tube.* Comprehension of the lower ranked words was within the grasp of kindergarten and first-grade children.

These same broad sampling considerations also were followed with the following 12 nonwords: *mōf, fūp, sim, ras, shig, thēk, hof, bīs, cheg, dak, tōb,* and *gās.* For each word type, half contained an initial sound that was continuant or could be held indefinitely, as in *rag* and *mof,* and half contained an initial nonresonate stop sound, as in *game* and *dak.*

A person knowledgeable in segmenting words into their component sounds recorded the 32 words using a Sharp cassette tape recorder that had excellent auditory quality and amplification (50/60 Hz, 24W). Each word was segmented at one of three intersound intervals, 0, 1, or 3 sec, entailing 96 presentations.

We presented the 32 words in 3 blocks. A word at one pause interval in one block did not appear at a different interval until the entire block of 32 words was exhausted. Words within each block were randomized and counterbalanced with the stipulation that no more than three pause intervals of the same duration would appear consecutively.

When recorded, each pronounceable sound was exaggerated, with continuant sounds said slowly and held for 2 sec and stop sounds said for a fraction of a second. For the three pause intervals, *not* was segmented as *nnoot, nn* (1-sec pause)... *oo* (1-sec pause) . . . *t.* The recorder was careful not to add the intrusive vowel sound to either held or stop sounds. Following each segmented word, the recorder said, "Say it fast," which was a pretaught direction to say the segmented word at its spoken rate.

Procedure. Prior to individualized testing on the target words, the children were taught to blend different word types through a "say-it-fast" game. Five-to-six children sat in a semicircle. The teacher modeled the game format and then tested first the entire group and, finally, individual children on their ability to blend different word types presented in this order: (a) two-syllable compound words with each part separated by a slight pause (e.g., *ice..cream, foot..ball*); (2) two-syllable simple words separated by a slight pause (e.g., *pen..cil, af..ter*); (3) a mix of CV (consonant-vowel) and VC (vowel-consonant) words and nonwords, with no pause between sounds (e.g., *mmēē, daa, aann, ūk*); and (4) a mix of CV and VC words with a 1-sec pause between sounds (e.g., *ss..ēē, nn..ūū, ii..ff, aa..t*).

To start the game, the teacher said, "I'll say a word slowly, then I'll say it fast. Listen: *ice..cream.* Say it fast..ice cream!" Then the group was given a turn with *ice cream* and the other compound words until everyone could blend each word without assistance. The children had no trouble blending the words from the first three word types, and only a few children had trouble with words from the fourth type. Although still a member of the group, each child was ultimately and individually tested in a random order with words from the fourth type until three consecutive correct blends were given. This test served as the criteria for participation in the study; only one child, who had unclear speech, could not be included.

One or two days following the familiarization-screening procedure, we began testing on the taped target words. The children were individually tested on the initial 48 words, then 2-3 days later on the last 48 words. Prior to testing, children were refreshed and, when necessary, coached on the say-it-fast game using VC and CV words. Noncontingent praise was intermittently given (e.g., "Good talking") during each of the 15-to 20-min testing sessions.

Interobserver agreement. We recorded all blended answers phonetically (e.g., *nose* = *noz*). during 18 randomly selected testing sessions, another trained observer independently scored the blending responses along with the main observer. The proportion of interobserver agreement was computed by the number of agreements (both observers agreed that the word was correctly or incorrectly blended) divided by agreements plus disagreements. The median agreement was .90 (range = .81 to 1.00).

Design. We used a 2 x 3 x 2 x 2 mixed ANOVA with fixed effects to analyze the data. The between-factor variable was age (kindergarten and first grade). The three within-factor variables were intersound interval (0, 1, and 3 sec), word familiarity (meaningful or nonword) and type of initial sound (held or stop).

Results

The means and standard deviations are included in Table 1. We found no significant four-way or three-way interactions. Four of the two-way interaction were significant, and the simple effects were evaluated by one-way ANOVAs. All reported *F* values were significant as $p < .004$ and t values as $p < .0001$.

The Interval x Age interaction was significant, $F(2, 84) = 17.99$. When further analyzed, first graders performed better than the kindergarteners at all three intervals: (a) 0 sec, $F(1, 42) = 10.06$; (b) 1 sec, $F(1, 42) = 29.22$; and (c) 3 sec, $F(1,42) = 47.83$. Within the kindergarten group, the interval main effect was significant, $F(2, 42) = 142.79$. Pairwise comparisons (using a modified Bonferonni t value of 4.89 based on Keppel, 1982) revealed better performance at 0 sec versus 1 sec, $t(21) = 11.92$, and 0 sec versus 3 sec, $t(21) = 13.77$. Differences between 1 and 3 sec were not reliable.

Within the first-grade group, the interval effect also was significant, $F(2, 42) = 50.79$, and pairwise comparisons yielded a similar pattern with performance better at 0 sec versus 1 sec, $t(21) = 7.89$ and 0 versus 3 sec, $t(21) = 7.74$. The 1- versus 3-sec differences were nonsignificant.

The Word Familiarity x Age interaction, $F(1, 42) = 11.58$, when further scrutinized, disclosed that first graders outperformed kindergarteners on both meaningful and nonwords: $F(1, 42) = 22.29$ and $F(1, 42) = 42.97$, respectively. Within the kindergarten group, the word familiarity main effect was significant, $F(1, 21) = 21.06$, with better performance evidenced on meaningful words. First graders, however, did not differ on the two word types, $F(1, 21) = 0.79$.

The significant interaction of Initial Sound x Word Familiarity, $F(1, 42) = 9.36$, revealed after further analysis better performances on both meaningful and nonwords with stop first sounds, $F(1, 42) = 9.72$ and $F(1, 42) = 26.48$, respectively. Analysis of word familiarity with held first sounds uncovered significantly more correct responses only on meaningful words, $F(1, 42) = 14.87$.

A breakdown of the significant Interval x Word Familiarity interaction, $F(2, 84) = 37.91$, uncovered: (a) at 0 sec, better performance on meaningful words, $F(1, 42) = 46.88$; (b) at 1 sec, better performance on nonwords, $F(1, 42) = 9.04$; and (c) at 3 sec, no significant differences between word familiarity. Concerning the significant differences of meaningful words, $F(2, 84) = 256.34$, pairwise comparisons (using a modified Bonferoni T value of 4.41) yielded better performance at 0 versus 1 sec, $t(43) = 14.69$; and 0 versus 3 sec, $t(43) = 16.39$. No differences occurred at 1 versus 3 sec. concerning the significant effect on nonwords, $F(2, 84) = 59.74$, pairwise comparisons also were significant for 0 versus 1 sec, $t(43) = 7.78$ and for 0 versus 3 sec, $t(43) = 7.13$, but not for 1 versus 3 sec.

Table 1. Percentage Correct Oral Blending Performance

	Kindergarten age		1st Grade age	
Condition	*M*	*SD*	*M*	*SD*
0–second pause				
Meaningful				
Stop 1st	75	21	82	11
Held 1st	64	18	77	22
Nonword				
Stop 1st	54	24	76	15
Held 1st	47	22	58	22
1–second pause				
Meaningful				
Stop 1st	19	23	46	23
Held 1st	13	18	41	26
Nonword				
Stop 1st	20	30	62	20
Held 1st	12	17	45	26
3–second pause				
Meaningful				
Stop 1st	15	16	50	24
Held 1st	17	20	42	20
Nonword				
Stop 1st	17	26	64	22
Held 1st	7	12	45	24

Note. Means and standard deviations are rounded to the nearest whole number.

Discussion

The likelihood that school-age children will have difficulty blending CVC words when the successive spoken sounds are broken by silent pauses is supported by the present findings and is consistent with their poor blending performance reported elsewhere (Chall et al. 1963; Williams, 1980). Although the 1- and 3-sec pause intervals produced diminished blending performance of equals magnitude, we are not certain whether the suppression is limited to durations of 1 sec or longer or whether any breaking sequence, however small, is sufficient. Because intersound intervals are not commonly specified in studies on blending, one needs a parametric investigation that included intervals of less than 1 sec to settle this issue. Interestingly, the short 0.5-sec delay reported by Chall et al. (1963) was accompanied by dismal blending performance.

The effect of pausing between sounds was much more deleterious for the kindergarten children, whose overall correct performance was 16% at 1 sec and 14 % at 3 sec, than for the first graders whose comparable performance was 49% and 50%, respectively. When the segmented sounds in a word were presented without any intervening pause, both age groups responded at much higher levels, although first graders still did better at 0 sec (73%) than the kindergarteners (60%). We will discuss the instructional implications of these findings later.

The differences between the two age groups could have resulted both from general experiential factors associated with the first graders' being a year older or specific factors associated with training in reading. We examined these possibilities in Experiment 2, where all children of kindergarten age (or younger) received training in reading through a program that taught them blending skills. Experiment 2 also enabled a replication of Experiment 1 findings.

It is unlikely that the reduced blending accuracy during pausing could be attributed to unfamiliarity with the task demands. The children appeared to understand the "say-it-fast" game, especially after extensive training in blending various word types without delays and in blending CV and VC words at a 1-sec delay. In addition, the segmented sounds were exaggerated and said slowly, in accord with Liberman's (1974) and Lewkowicz's (1980) conclusions that a "stretched" pronunciation of the word should help the child perceive the separated sounds.

On the other hand, presenting the target words through taped recordings, as is done in standardized blending tests (Goldman, Fristoe, & Woodcock, 1976; McCarthy & Kirk, 1961), may have contributed to a poorer performance than that achieved if the segmented sounds were said in full view of the child and articulatory movements of the mouth served as additional cues.

Some finding relating to the effects of word familiarity and type of initial sound are not clear. As expected, the kindergarten children performed better on the meaningful versus the nonwords, whereas the first graders unexpectedly did not differ on word familiarity. By being better all-around blenders, the first graders did not need to rely on the semantic value of the word as much as the kindergarten children did.

The findings that beginning stop-sound words led to better performance than those with beginning held sounds (true for both meaningful and nonwords) was unexpected and incongruent with previous research (Coleman, 1970; and Williams, 1980). Because beginning held sounds lasted longer, they should have aided the recall of sounds more than the split-second stop sounds did. Stop sounds, by supposedly causing the learner to add an extra vowel during the recombining process, should have led to greater word distortions (Gleitman & Rozin, 1973). The narrator, while making the tape, intently avoided the addition of an intrusive vowel. Two possible accounts for the advantage of stop sounds follow: (a) By being sharp and brief, stop sounds heighten attention to themselves and to the ensuing sounds; (b) Stop sounds reduce the total duration of time spent in synthesizing words that contain them and result in less demand on short-term auditory storage.

The significant Intersound Interval x Word Familiarity interaction revealed that meaningful words were blended better at 0 sec, whereas nonwords were blended better at 1 sec (no differences appeared at 3 sec). The surprising differences at 1 sec were due largely to the vast differences between the first graders and the kindergarten children at this delay interval.

Table 2. Last Reading Lesson Completed and Ages for Reading Groups.

Group		Last Lesson	Age[a]
Advanced	*M*	274	71.0
	SD	25	9.6
Intermediate	*M*	139	64.8
	SD	31	12.2
Beginning	*M*	32	51.9
	SD	29	8.1

Note. Maximum number of reading lessons = 320.
[a] Ages are in months.

Experiment 2

Method

Subjects. The children attended a preschool that used the first two levels of the SRA Reading Mastery program (Engelmann & Bruner, 1983). Each level contained 160 lessons. The children were classified according to the last reading lesson completed at the time of the blending test. Three groups ($n = 17$ each) emerged: advanced, intermediate, and beginning.

According to Beck and McCaslin (1978), Reading Mastery is a code-emphasis program that provides a definite instructional strategy for teaching blending. Oral blending is one prerequisite skill for reading, and, when taught, no pauses occur between sounds.

For the three training groups, there were no group differences on such factors as racial composition, sex, and proportion of children eligible for free meals. There also were no initial Slosson IQ differences.

Table 2 gives the group breakdown by last lesson and age. The groups differed on number of lessons completed, $F(2, 48) = 310.97$, $p < .0001$ with all intergroup comparisons (Bonferroni t) reliable, all $ps < .001$. The groups also differed according to age, $F(2, 48) = 15.98$, $p < .0001$, with the children in the advanced and intermediate groups each significantly older than the beginning group, $ps < .01$. The advanced versus intermediate age differences fell short of significance, $p = .08$.

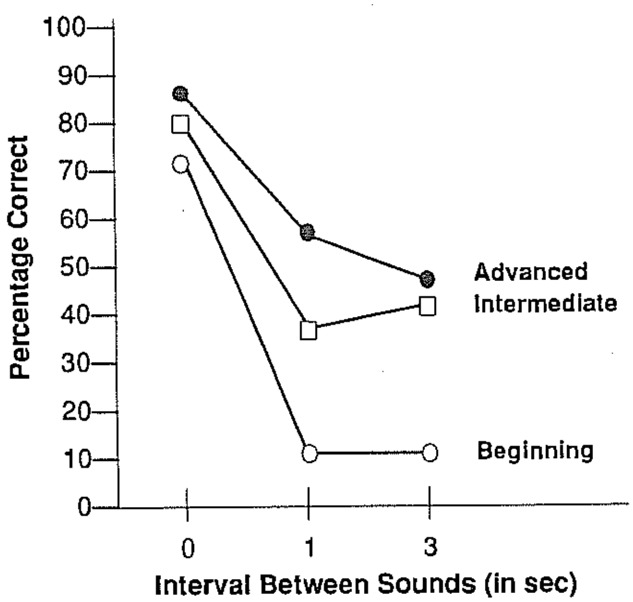
Procedure. In Experiment 2 we followed the same familiarization and testing procedures as those in Experiment 1.

The median interobserver agreement, based on two observers independently scoring 27 children, was .96 (range = .88 to 1.00).

Results

There were no significant four-way or three-way interactions. The Reading Group x Interval interaction, the nature of which appears in Figure 1, was significant, $F(4, 96) = 12.92, p < .0001$. Three separate, one-way ANOVAs yielded reliable between-group effects at each interval, all $Fs(2, 42) \geq 5.51$ and all $ps <$ 007. Pairwise comparisons were completed using modified Bonferroni t values of 2.19, 3.62, and 4.37 for significance levels of .05, .001, and .0001, respectively. The advanced and intermediate groups did not differ at any interval, all $ts(32) < 2.16, ps > .05$. Both groups did significantly better than the beginning group at all intervals, with the differences being more pronounced at the 1-sec and 3-sec intervals, $ts(32) > 4.08, ps < .001$, than at 0-sec, $ts(32) > 2.50, p < .05$. Performance within interval levels revealed that each group blended significantly better at 0 sec versus 1 sec, all $ts(16) > 8.71$, all $ps < .0001$, and at 0 sec versus 3-sec, all $ts(16) > 8.51$, all $ps < .0001$. The 1-sec versus 3-sec differences were not significant for any reading group.

Figure 1. Mean Percentage Correct Blending Performance for the Reading Group x Interval Interaction in Experiment 2.



As in Experiment 1, significant interaction occurred between interval and word familiarity, $F(2, 96) = 5.60$, $p < .005$, and between initial sound and word familiarity, $F(1, 48) = 6.89, p < .01$. an analysis of the variables contained in these interaction yielded essentially the same pattern as that described in Experiment 1. Unlike the findings in Experiment 1 of a significant Word Familiarity x Age interaction, the analogous interaction for Experiment 2 (i.e., Word Familiarity x Reading Group) fell short of significance, $F(2, 48) = 2.73, p = .08$.

Blending performance correlated .69 with the number of reading lessons, .40 with age, but only .11 with entry IQ. To statistically control for the effect of age and to assess the contribution of the number of reading lessons on blending performance, we conducted an ANOCOVA with age as the covariate. Results indicated that the reading group main effect previously revealed by the ANOVA was upheld, $F(2, 47) =$ 13.80, $p < .0001$; covariate, $F(1, 47) = .16, p = .69$.

A stepwise multiple regression analysis showed that the number of lessons was the best significant predictor, $t(50) = 5.49, p < .01$, accounting for over 46% of the blending variance. The next best predictor was IQ, $t(50) = 2.88, p < .01$, but it contributed only an additional 7 % to the multiple R-square. Age entered the full-model analysis last and was not a significant predictor, $t(50) = 1.11, p = .27$, contributing merely 1% to the total.

General Discussion

The apparent justification for injecting pauses between individual or clusters of sounds during oral blending is to demonstrate that words are divisible into parts. If pausing serves this important function, then it is more than offset by preventing the naive learner from recombining these broken parts into whole words with a high degree of accuracy (at least with the pause intervals used herein). The data suggest that the teacher say the sounds slowly in an exaggerated form without introducing any pauses between them. To help convey that words consist of smaller units, the teacher could emphasize changes in sound values by exaggerating mouth movements or by clapping or holding up a finger for each new sound said.

New learners also would benefit by initially blending sounds that form familiar words, and as this skill becomes perfected, sounds from unfamiliar words (or nonwords) would then be programmed. For both kinds of words, however, no intersound pauses should occur.

Improved accuracy in blending words orally, because a teacher (or narrator) said the sounds in a continuous manner, appears to prepare students for decoding printed words (Carnine & Silbert, 1979). If students can be taught to sound out words without pausing between the sounds as the teacher did during oral blending, the process of decoding words should be smoother and more accurate than if the student paused or hesitated between sounds.

Occasionally, however, pausing is necessary. When spelling dictated words, the student is likely to pause momentarily as the word is segmented orally into its parts. The student then probably engages in longer pauses as each pronounced part, either the sound value or the alphabet name of each letter, is converted into its written form to complete the word. Because of their spelling training, the first graders in Experiment 1 and the kindergarten children in the advanced and intermediate groups in Experiment 2 probably learned much about how to derive words under conditions of self-imposed pauses. During oral blending at the 1- and 3-sec pauses, we observed many of these children making nonvocal lip and mouth movements after each dictated sound. Then, following the last sound, the children attempted to repeat all the sounds in sequence prior to saying the word. A distinct possibility exist that the children sometimes arrived at the word by spelling it phonetically.

The kindergarten-aged advanced and intermediate readers blended as well as the first graders across all pause intervals. On the other hand, blending by the nonreading kindergarten children of Experiment 1 was at substantially lower levels, especially when pauses were imposed, and was almost identical to the younger-aged beginning readers. These findings together show that oral blending can be enhanced through direct instruction is not necessarily one of the phonemic awareness deficits associated with kindergarten-aged children (Calfree, Lindamood, & Lindamood, 1973; Rosner & Simon, 1971).

The results should not imply that blending training itself was responsible for the changes in oral blending, because we did not employ an independent group receiving instruction in oral blending but not in reading, as other researchers did (Goldstein, 1976; Rohnback, Bell, & McLaughlin, 1982). The fact that the number of reading lesson correlated highest with an ability for blending and accounted for the largest proportion of blending variance is indirect evidence for the promotion of oral blending by reading-related experiences. Advocates of the whole-language approach (Altwergen, Edelsky, & Flores, 1987) might want to expand upon this point: One spin-off of the emerging literacy of young reader could be in an improvement in their ability to blend words. ❖

Acknowledgments

Support from the University of Alabama's Office of Sponsored Programs is gratefully acknowledged. We also extend our appreciation to Marc Hendler, Karla Johnson, Nicholle O'Dell, and Connie Sulentic for their invaluable help in various phases of this project.

References

Altwergen, B., Edelsky, C., & Flores, B.M. (1987). Whole language: What's new? *The Reading Teacher, 41*(2). 144–154.

Barnard, D.P., & DeGracie,J. (1976). Vocabulary analysis of new primary reading series. *The Reading Teacher, 30*, 177–180.

Beck, I. L., & McCaslin, E. S. (1987). *An analysis of dimensions that affect the development of code-breaking ability in eight beginning reading programs.* Pittsburgh: University of Pittsburgh, Learning Development Center.

Calfee, R. C., Lindamood, P., & Lindamood, C. (1973). Acoustic-phonetic skills and reading: Kindergarten through twelfth grade. *Journal of Educational Psychology, 64*, 293–298.

Carnine, D.W., & Silbert, J. (1979). *Direct instruction reading.* Columbus, OH: Merrill.

Chall, J., Roswell, F. G., & Blumenthal, S. H. (1963). Auditory blending ability: A factor in success in beginning reading. *Reading teacher, 17*, 113–118.

Coleman, E. B. (1970). Collecting a data base for a reading technology. *Journal of Educational Psychology Monograph, 61*(4), 1–23.

Dolch, E. W. (1963). A basic sight vocabulary. *Elementary School Journal, 36* 456–460.

Engelmann, S., & Bruner, E. (1983). *Reading Master I and II, DISTAR Reading.* Chicago: Science Research Associates.

Gleitman, L. R., & Rozin, P. (1973). Teaching reading by use of a syllabary. *Reading Research Quarterly, 8*, 447–483.

Goldman, R., Fristoe, M., & Woodcock, R. W. (1976). *Goldman-Fristoe-Woodcock Auditory Skills Test Battery.* Circle Pines, MN: Americian Guidance Service.

Goldstein, D. M. (1976). Cognitive-linguistic functioning and learning to read in preschoolers. *Journal of Educational Psychology, 68*, 680–688.

Haddock, M. (1976). Effects of an auditory and an auditory-visual method of blending instruction on the ability of prereaders to decode synthetic words. *Journal of Educational Psychology, 68*, 825–831.

Katz, J., & Harmorn, C. (1982). *Phonetic synthesis: Blending sounds into words.* Allen, TX: Developmental Learning Materials.

Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd.). Englewood Cliffs, NJ: Prentice-Hall.

Lewkowicz, N.K. (1980). Phonemic awareness training: What to teach and how to teach it. *Journal of Educational Psychology, 72*, 686–700.

Liberman, A. M. (1974). The specialization of the language hemisphere. In F. O. Schmitt and F. G. Woodsen (eds.), *The neurosciences: Third study program.* Cambridge, MA: MIT Press.

McCarthy, J.J. & Kirk, S. A. (1961). *Examiner's manual: Illinois Test of Psycholinguistic Abilities (Experimental version.)* Urbana, IL: University of Illinois Institute for Research on Exceptional Children.

Muller, D. (1972–1973). Phonic blending and transfer of letter training to word reading in children. *Journal of Reading Behavior, 5*, 212-217.

Richardson, E., DiBenedetto B., & Bradley, C. M. (1977). The relationship of sound blending to reading achievement. *Review of Educational Research, 47*, 319-334.

Rohnback, C. R., Bell, B. J., & McLaughlin, T. F. (1982). The value of auditory blending skills for reading. *Educational Research Quarterly, 7*, 41-47.

Rosner, J. (1973-1974). Auditory analysis with prereaders. *The Reading Teacher, 27*, 379–384.

Rosner, J., & Simon, D. P. (1971). The auditory analysis test: an initial report. *Journal of Learning Disabilities, 4*, 384–392.

Roswell, F. G., & Chall, J. (1963). *Auditory blending test.* New York: Essay Press.

Weisberg, P. (in press). DI in the preschool. *Education and Treatment of Children.*

Williams, J. P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology, 72*, 1–15.

Using Choral Responding to Increase Active Student Response*

by William Heward
Frances H. Courson
Ohio State University, Columbus, Ohio
Janani Marayan
Fresno Unified School District, Fresno, California

More than 70 years ago, Dewey (1916) emphasized that students learn by doing. But educational researchers have only recently rediscovered this essential variable in the teaching-learning formula. Investigators who use the research methods of group comparison/statistical inference and their colleagues who favor repeated measures/single-subject analysis have followed strikingly similar paths in the past decade and a half. Both groups have consistently found the same relationship in a long and still-growing series of studies involving learners of all ages and characteristics.

According to the group comparison researchers, academic achievement is more likely to occur in classrooms in which students are actively engaged with instructional materials on which they have a high success rate. This student engagement is described as *academic learning time*, or *ALT*. (For a comprehensive review of ALT literature, see Fisher & Berliner, 1985.)

According to the behavior analysts, a functional relationship exists between academic performance and the frequency of active responses students make during instruction. The construct *opportunity to respond (OTR)* has been used to measure active student response in the classroom (Stanley & Greenwood, 1983). Greenwood, Delquadri, and Hall (1984) have provided an excellent review of the research showing that increased OTR results in improved student achievement.

It appears the Dewey was right all along. But how does one teacher with many students arrange sufficient opportunities for all of them to respond? Fortunately, there are numerous practical strategies for increasing active student response during group instruction. One of these strategies, choral responding, is the subject of this article.

*From Using Choral Responding to Increase Active Students Response by Heward, Courson and Narayan, *Teaching Exceptional Children*, Spring, 1989, 72-75. Copyrighted 1989 by The Council for Exceptional Children. Reprinted with permission.

Choral Responding

Choral responding (CR)—all students in the group orally responding in unison to a teacher-posed question—is certainly not a new idea. But its contemporary use is limited. It is true that children in primary classrooms are occasionally asked to respond in unison, and CR is a common instructional method of foreign language teachers. However, it is employed systematically as a daily teaching routine by very few teachers today. The procedure is ideally suited for increasing the frequency of active student response during group instruction, and we believe that teachers should use it more often.

Research Background

Sindelar, Bursuck, and Halle (1986) compared the effectiveness of ordered questioning with unison responding (CR) in teaching sight words to 11 elementary school students classified as having learning disabilities or mild mental retardation. Instruction took place in groups of 3 or 4 students. When presenting words via ordered questioning, the teacher called upon individual students in a left-to-right sequence around the group. When words were taught with the unison responding method, all students in the group responded together at the signal "Everybody." Results showed a small but significant advantage on favor of unison responding in both the rate of learning new sight words and the percentage of previously learned words correctly identified on a posttest.

McKenzie and Henry (1979) taught a standard science lesson to 52 third graders. They compared a variation of CR, in which students responded in unison with an "interpretable nonverbal gesture" (e.g., raised hand, pointing to an object) to a procedure in which individual students were called upon to respond. Students in the individual-response group were off task twice as often as those taught with the in-unison procedure. Achievement, as measured by a 30-item posttest, was also better for the students in the in-unison response group.

Although few research studies have investigated CR as a teaching procedure in its own right, CR is a major component of DISTAR (Engelmann & Bruner, 1974), one of the most widely studied and effective instructional programs ever developed. DISTAR consists of a scripted series of lessons to which small groups of children respond both individually and chorally to a fast-paced series of teacher-presented questions, prompts, and cues. The nationwide evaluation of learning outcomes of low-income minority students in Project Follow Through showed that children who have participated in programs using the Direct Instruction Model (DISTAR) made more achievement gains in reading and mathematics than did children who had participated in programs based of any of nine other models (e.g., Piagetian, parent-centered, discovery learning) (Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977). This 9-year study involved more than 8,000 students in 180 school districts. For a comprehensive review of the research on DISTAR's effectiveness see Gersten, Carnine, and White (1984).

When to Use CR

In most cases, CR sessions should be limited to 5 to 10 minutes at a time. CR is fun for both teacher and students, but it demands everyone's complete attention and active involvement. It is best employed by having students participate in several short CR sessions for different subjects throughout the school day.

CR can be done with any size group, but it is perfectly suited to the small groups with whom special educators often work. It is also an excellent technique for actively engaging mainstreamed students in regular classroom instruction. CR is used most appropriately with curriculum materials that meet the following criteria.

1. *Short answers.* Students must be able to respond briefly to the questions or items presented during CR. Identifying parts of speech, giving answers to mathematics problems, voicing the sound of consonant blends, and stating the parts of green plants having different functions are just a few examples of curriculum areas in which CR can be used. Questions requiring students responses that are longer than one to three words destroy the rhythm and pacing that are critical to effective CR. Long student responses also make it difficult for the teacher to provide feedback effectively because only a jumbled mix of voices can be heard.

2. *Only one correct answer.* There should be one correct answer for each item presented during CR. The teacher must listen for a single correct response and provide feedback to the group accordingly. Hearing multiple correct responses and discriminating those from any accompanying incorrect responses given to the same question would be too difficult fir the teacher, and the teacher's feedback would be too complicated for students to comprehend. Clearly, this requirement limits the use of CR. It is simply inappropriate for curriculum content for which divergent, creative responses are desired.

3. *Presented at a fast pace.* CR is most effective when conducted at a lively pace. The relationship between teacher pacing of CR and student performance was demonstrated in a study by Carnine (1976). Four low-achieving first graders received 30 minutes of DISTAR instruction each day. The study entailed two conditions: a slow pace, in which the teacher paused for 5 seconds between providing feedback for the student's response to one question and presenting the next question and presenting the next question, and a fast pace, in which the teacher presented the next item immediately after providing feedback for the previous trail's responses. Data were taken on the students' off-task behavior, participation (percentage of questions to which each child responded), and accuracy of responses during instruction. Results showed a dramatic difference in favor of the fast-paced presentation.

A recent systematic replication of Carnine's study was conducted with 8- to 12-year-old students with severe behavior disorders using a 3-second intertrial interval for the slow presentation pace (Morgan, 1987). The same pattern of results was obtained in favor of the fast-paced presentation.

When deciding whether or not a lesson would benefit from CR, the teacher should determine how long it takes to present each question or item for student response. If presentation takes too long, the pacing and rhythm of CR will be lost and too few opportunities for active student provided. A minimum rate of about 4 student responses per minute should be maintained. Response rates of up to 10 per minute can be achieved as students develop good skills with material such as sight words and mathematics facts.

How to Conduct CR

Choral responding is most effective and most enjoyable if it is conducted in a systematic and energetic manner. The following procedures and guidelines are recommended:

1. *Model the response.* Students must understand what kind of response they are to make. Begin the CR session by modeling the type of response desired. For example: "I'm going to read some short story problems to you when I signal 'What operation?' answer by saying either 'addition,' subtraction,' multiplication,' or 'division.' Listen and I'll show you how to do it." Then model several trials, playing both your own and the students' roles.

2. *Provide a "thinking pause."* Depending on the complexity of the material and the students' level of mastery, a "thinking pause" might be necessary. This pause—a few seconds between asking the question and signaling students to respond—gives students an opportunity to figure out their answer. The thinking pause should not be too long; 4 to 6 seconds are appropriate for many activities. Questions requiring

more than 10 seconds to answer are not well suited to CR because they interfere with the necessary rhythm of the session.

As students' skills develop there are some CR activities for which an immediate response is desirable. Sight words and mathematics facts, for example, are areas in which a proficient rate of correct responding is important. If a student solves 100% of the addition and subtraction facts correctly but is working at a rate of only 5 problems per minute, the student still has a long way to go to mastery. By gradually reducing the thinking pause during CR, you can help your students become more proficient.

3. *Provide a clear signal for response.* Choral responding works best when all students in the group respond in unison. Teach your students to respond immediately to a standard signal. You might use a word such as "Students," "now," or "Class" as your signal, or you might prefer a physical cue such as dropping your arm or pointing to the question on the blackboard or overhead projector. Whatever signal you choose, practice it with your students so they develop a good pattern of unison responding.

When providing a thinking pause of longer than 3 seconds, say "Get ready!" immediately before giving the response signal. This cue greatly improves the unity of the group's response because it allows students who have determined their answer several seconds earlier to get ready to say it on cue.

4. *Provide feedback for the "majority" response.* Choral responding consists of a fast-paced series of three-part learning trials: teacher-presented questions, student response, and teacher-delivered feedback. Give immediate feedback for every response from the group. When you hear only correct responses a quick "Yes!" or "Great" is sufficient.

When one or a few incorrect answers are heard, repeat the correct answer before presenting the next item. For example, "Yes, multiplication is the correct operation for solving that problem." When a third or more of the group gives the incorrect answer, repeat the correct answer and then immediately repeat the item for choral response. Several trials later, repeat questions to which any incorrect answers were given.

5. *Call upon individual students occasionally.* Although students respond as a group in CR, your primary concern is, as always, each child's progress. Two procedures can be used to assess the performance of individual students during CR. The first involves simply listening for the voice of the student in question. Directing your eye contact along with a special smile to a student who has just responded correctly can be very reinforcing. When you hear an incorrect answer given by a student, pay particular attention to his or her response when you repeat the same question several trials later.

The second method for assessing individual performance is accomplished by mixing requests for individual student responses with requests for group responses. Be sure to ask the question before you call an individual student's name so that you do not inadvertently signal the rest of the group not to pay attention to that particular item. By asking the question first, you alert each student in the group to prepare a response covertly in case he or she is called on individually or as a member of the group.

To increase attention to the lesson and motivate students to respond at the appropriate time, you might implement a group contingency system. The whole class could receive a reward for responding together as well as for making less than some predetermined number of choral responses when an individual student's name was called.

Some teachers are concerned that a student who is having difficulty with the material might just "follow along" during CR, voicing the correct responses only after hearing the other group members respond. We view such imitation as an advantage of CR, not a disadvantage, for children can benefit from peer models of correct responses. Calling upon such a student from time to time to check his or her ability to respond independently is a good way to determine whether or not such benefits are occurring. However, individual students should not be called on in an effort to catch them in an error. The inability to give a correct answer in front of their peers is an all-too-common experience for many students with learning problems. Frequent use of the individual-response tactic can provide these students with an opportunity to be successful in front of others. Regular classroom teachers should use this tactic often when conducting CR with mainstreamed children.

While we could find no research demonstrating an optimal percentage of group to individual responses, approximating the 7:3 ratio of choral to individual responses recommended by Stevens and Rosenshine (1981) is probably good practice.

6. *Maintain a lively pace.* The primary rationale behind CR is the large number of active learning trials it provides every student in the group. Quickly presenting one item after another gives students more opportunities to respond during the session than does a slower-paced presentation. Students are also more likely to stay one task, respond more often, and re-

spond more accurately when the teacher keeps a lively pace (Carnine, 1976; Morgan, 1987).

However, maintaining a lively pace does *not* mean rushing your presentation of each question, giving students an insufficient thinking pause, or providing insufficient feedback. Proper pacing is primarily a function of the following four variables: (1) choosing material that can be presented as a series of questions or items, each one of which can be stated or displayed within a maximum of 10 seconds; (2) providing a thinking pause just long enough for students to get ready to respond; (3) giving direct and immediate feedback; and (4) presenting the next question immediately after completing feedback for the previous trial.

While the first three variables will, in fact, vary with the subject matter being covered and the students' skill level, the fourth factor—intertrial interval duration—should always be as short as possible. Presenting the next item immediately after giving feedback for the previous question requires preparation. While a fast-paced presentation of a long series of items for some content areas (e.g., mathematics facts) can be ad-libbed," conducting a good CR session for most subject matter requires preparation. Using an overhead projector to display each successive item on a transparency is an excellent way to maintain a good pace while concentrating on your students' performance (Cooke & Test, 1984).

Remember that the optimal rate of learning trials per minute will vary considerably, from 4 to as many as 10, depending on the material and your students' skill level. Keep some data on your presentation rate, and with a little practice you will become an expert at managing CR.

Variations

When you and your students are accustomed to using CR, you will probably want to experiment with variations on the standard procedure. One variation is to let students try the role of teacher from time to time. Depending on the complexity of the material and the students' skills, you might prepare a teaching script of transparency with the correct answers following each question or let the new "teacher" provide feedback on his or her own. Playing the teacher role provides excellent practice on the material and is a special job that many students will work hard to earn.

Another variation is to replace oral responses with easy-to-see hand or finger responses (McKenzie & Henry, 1979). In one study, students in 10 fifth-grade classrooms signaled their solutions to mathematics problems by holding up the number of fingers corresponding to their choice of answers (Pratton & Hales, 1986). Students who were instructed with this active-participation method performed better on a posttest than students taught in a nonactive participation format.

Conclusion

Choral responding is an easy, virtually cost-free method of increasing each student's frequency of active response during group instruction. When it is properly conducted, both you and your students will enjoy CR. Remember the two basic rules: Do not conduct CR sessions for more than 5 to 10 minutes and keep up that pace! ◆

References

Carnine, D. W. (1976). Effects of two teacher presentation rates on off-task behavior, answering correctly, and participation. *Journal of Applied Behavior Analysis*, 9(2), 199-206.

Cooke, N. L. & Test, D. W. (1984). Instructional transparencies for the directive teacher. *The Directive Teacher*, 6(1), 22-23.

Dewey, J. (1916). *Democracy and education.* New York: Macmillan.

Engelmann, S. & Bruner, E. (1974). *DISTAR reading level 1.* Chicago, IL: Science Research Associates.

Fisher, C. W. & Berliner, D. C. (eds.). (1985). *Perspectives on instructional time.* New York: Longman.

Gersten, R., Carnine, D. W., & White, W. A. T. (1984). The Pursuit of clarity: Direct instruction and applied behavior analysis. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter, Eds.), *Focus on behavior analysis in education* (pp. 38-57). Columbus, OH: Merrill.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1984). Opportunity to respond and student academic achievement. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter, Eds.), *Focus on behavior analysis in education* (pp. 58-88). Columbus, OH: Merrill.

McKenzie, G. R. & Henry, M. (1979). Effects of testlike events on on-task behavior, test anxiety, and achievement in a classroom rule-learning task. *Journal of Educational Psychology*, 71, 370-374

Morgan, D. J. (1987). *Effects of fast and slow teacher presentation rates on the academic performance of special education students during small-group reading instruction.* Unpublished master's thesis, The Ohio State University, Columbus.

Pratton, J. & Hales, L. W. (1986). The effects of active participation on student learning. *Journal of Educational Research*, 79(4), 210-215

Sindelar, P. T., Bursuck, W. D., & Halle, J. W. (1986). The effects of two variations of teacher questioning on student performance. *Education and Treatment of Children*, 9(1), 56-66.

Stanley, S. O. & Greenwood, C. R. (1983). How much "opportunity to respond" does the minority disadvantaged student receive in school? *Exceptional Children*, 49, 370-373.

Stebbins, L., St. Pierre, R.G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). *Education as experimentation: A planned variation model* (Vols. A-D). Cambridge, MA: Abt Associates.

Stevens, R. & Rosenshine, B. (1981). Advances in research and teaching. *Exceptional Education Quarterly*, 2, 1-9.

Applications of Peer Tutoring to Arithmetic and Spelling

by Gregory F. Harper
Barbara Mallette
Larry Maheady
SUNY–Fredonia, NY
Rachel Clifton
Fredonia Central School Disttict

Classwide Peer Tutoring (CWPT) is a peer-mediated, instructional intervention developed by researchers at Juniper Gardens Childrens' Project in Kansas. (For a complete description, see Delquadri, Greenwood, Stretton & Hall, 1983; Delquadri, Greenwood, Whorton, Carta & Hall, 1986; Greenwood, Delquadri & Carta, 1988.) CWPT was designed to improve the basic skill performance of low-achieving minority, disadvantaged, and/or mildly handicapped students. CWPT consists of four basic components: (1) weekly competing teams, (2) highly structured teaching procedures, (3) daily point earning and public posting of pupil performance, and (4) direct practice of functional academic skills. Each week, the class is divided randomly into two competing teams. The teacher then assigns students *within each team* to tutoring pairs. One student in each pair serves as tutor for 10–15 minutes, while the other student is the tutee. After the established time limits have elapsed, the tutoring pairs reverse roles for an equivalent length of time.

While students work in their tutoring dyads, they follow prescribed instructional procedures. That is, the tutors present instructional items, (e.g., spelling words, math problems, social studies questions) and the tutees must "say and write" their responses. If an answer is correct, the tutor awards two points. However, if tutees responds incorrectly, the tutor: (1) models the correct response, (2) requires the tutee to write the answer three times, and (3) gives one point if the tutee corrects the mistake. If the tutee fails or refuses to correct errors, no points are awarded. Students are encouraged to complete as many items as possible within the time limit. The more items completed, the more points they earn for themselves and their team.

While students are tutoring one another, the teacher moves about the classroom awarding "bonus points" for good tutor and tutee behaviors. Immediately after the tutoring sessions, students total their daily points, including bonus points, and record them on a laminated scoreboard posted in the classroom. Tutoring sessions are conducted three to four times per week and are followed by a weekly test. Tests are taken *individually*; however, students receive five team points for each correct test item. At the end of the week all points, including test points, are totaled and the "Winning Team of the Week" is announced. Weekly results, as well as outstanding individual performances, are then officially recognized either in classroom/school bulletin or via achievement certificates.

To date, the academic effects of CWPT have been documented extensively at the elementary school level by Juniper Gardens staff (e.g., Delquadri et al., 1983; Greenwood, Delquadri & Hall, 1989; Greenwood, Dinwiddie, Terry, Wade, Stanley, Thibadeau & Delquadri, 1984), as well as independent evaluators (e.g., Cooke, Heron & Heward, 1983; Maheady & Harper, 1987). Data from these studies have shown CWPT to be powerful intervention for improving students' reading, math, and spelling skills. Recently, CWPT procedure were extended to the secondary level where significant improvements were also obtained with mildly handicapped students enrolled in social studies and math classes (Maheady, Harper & Sacca, 1988; Maheady, Sacca & Harper, 1987; Maheady, Sacca & harper, 1988).

The reasons for the success fo CWPT can be attributed to its impact on the allocation of time to instruction and on opportunity to respond (Greenwood, Delquadri, & Hall, 1984). CWPT requires active involvement with learning tasks, permits immediate feedback (both positive and corrective) and allows the learner to practice in a context which is topographically similar to the criterial task. Hence, for example, the stimulus (orally presented spelling words, oral reading assignments) and the student response (written words, oral reading) are the same in tutoring as in testing or evaluation (e.g., spelling tests, group oral reading sessions). CWPT is therefore highly consistent with the tenets of direct instruction.

In this paper, we briefly report three demonstrations of CWPT undertaken at the request of classroom teachers. In each instance we were responding to perceived instructional problems. Largely as a result, these demonstrations lack certain elements of a rigorous experimental design. Results reflect realistic classroom conditions and constraints and are perhaps best thought of as naturalistic case studies. Procedures in each demonstration follow those described above, with the exceptions noted.

* Complete manuscripts describing each of these studies are available from the senior author.

I. Use of CWPT to Increase the Fluency of Mathematical Computations Among Second Graders.

District-wide mandates, based on interpretations of state-wide curricular standards, required students to complete 100 mathematics problems with 90% accuracy in 12 minutes. In this instance, single digit subtraction problems with minuends of less than 16 were targeted. Although many students "knew" the solutions—pretest accuracies indicated a class average of 95.7%—fewer than half of the students met the rate and accuracy criteria. Four students required more than 20 minutes to complete the 100 problems. Prior teacher efforts to increase rate, and to maintain or improve accuracy had included timed tests, the use of flashcards with individual students and homework practice sheets. Increased rates of responding are a common outcome of CWPT procedures. In this case we wondered if it could be used to increase rate as a primary target, while maintaining accuracy and reducing variability among students.

Subjects

Subjects were 17 children (9M, *F) enrolled in a second grade classroom in rural western New York State (mean age = 107.8 months, s.d. = 6.68, range 96–119 months).

Procedures

CWPT was inplemented as described elsewhere (Greenwood, et al., 1986). Materials consisted of 16 or 17 problems written on cards. Each dyad had its own set of cards. Use of cards permitted randomized presentation of problems. The tutor read the problem, the tutee wrote and simultaneously said the problem and the solution. New problems were presented each wek from a pool of 100. CWPT continued for four days each week for 10 weeks.

Results

Outcome measures included rate and accuracy tests each Friday on those problems practiced that week. a post-test over the 100 problems at the end of 10 weeks was administered. Weekly accuracy rates and rates of responding were calculated by counting the total number of times each problem was attempted by each student on a weekly basis. Results are presented in Table 1.

Accuracy for the 10-week period varied between 91.6% and 98.9% for the entire class. For individual students the lowest accuracy in any week was 70%, the highest 100%. The average number of seconds per problem on weekly tests varied between 4.29 and 5.83 seconds, with a mean of 4.98 seconds. The gain on the 100 item test was from 95.7% on the pretest of 98.1% on the posttest.

Individual student scores varied between 93% and 100% on the posttest. The average time for computation of 100 problem dropped from 723 seconds on pretest to 546.11 seconds on the posttest. All students improved their time with 15 of 17 (or 88%) meeting the 12 minute, 90% criteria. Only 47% met it on the pretest.

Discussion

The results suggest that CWPT can be used as a means to improve rate of responding while maintaining high accuracy. The student performances was improved on both accuracy and time outcomes. Stu-

Table 1. Percent Correct and Times on Pre- and Posttests and Weekly Quizzes Class Average

	Pretest	Week Number										Posttest
		1	2	3	4	5	6	7	8	9	10	
Percent Correct	95.7	97.8	95.1	97.8	94.5	96.6	98.4	97.2	95.4	98.9	97.5	98.1
Standard Deviation	8.29	3.71	5.71	3.72	9.50	4.36	2.75	4.79	5.42	2.41	3.71	1.98
Mean Time (seconds)	723	72.9	89.4	82.2	93.4	73.3	78.1	77.2	94.2	73.3	93.3	546
Number of Items	100	17	17	17	16	16	16	16	16	16	16	100
Mean Time per Item (seconds)	7.23	4.29	5.25	4.83	5.50	4.58	4.88	4.82	5.26	4.58	5.83	5.46
Percentage Meeting Criterion	47											88

dent ratings of the CWPT procedure indicate high levels of consumer satisfaction (i.e., "They liked it").

II. Generalization of Spelling Words.

Although previous research has indicated that the use of CWPT improves student spelling performance on weekly spelling tests (Maheady & Harper, 1987; Delquadri, Greenwood, Stretton & Hall, 1983), there is little evidence to date that these gains are generalized to tasks outside of practice and testing. One would hope that words spelled correctly on Friday tests *would be used correctly in other applied settings.* To assess the effect of CWPT on this generalization, students were assessed on their ability to write dictated sentences containing the spelling words from previous weeks.

Subjects

Seventeen children (9F, 8M) enrolled in a second grade classroom in a rural school district in western New York state served as subjects. Their mean age was 89.6 months (s.d. = 5.07, range 82–99 months).

Procedures

An AB design was used in this study. Three weeks of baseline data were taken reflecting student performance under weekly spelling instruction (i.e., the procedures generally used during spelling lessons). A maintenance probe was implemented in week two of baseline and for all remaining 15 weeks of the study. In each instance, five sentences were dictated by the classroom teacher, containing 10 words from previous weeks. In Week 2, all 10 words were from Week 1; thereafter (Weeks 3–15) the words contained in the sentences included six words randomly selected from the previous week's list, and four words randomly selected from all prior weeks. for example, the sentences transcribed by students in Week 3 contained six words from Week 2 and four words form Week 1.

During the fourth week, a CWPT procedure was implemented (Maheady & Harper, 1987). Procedures described above were utilized by students to practice 10 spelling words. the words were those contained within the regular spelling curriculum for this class.

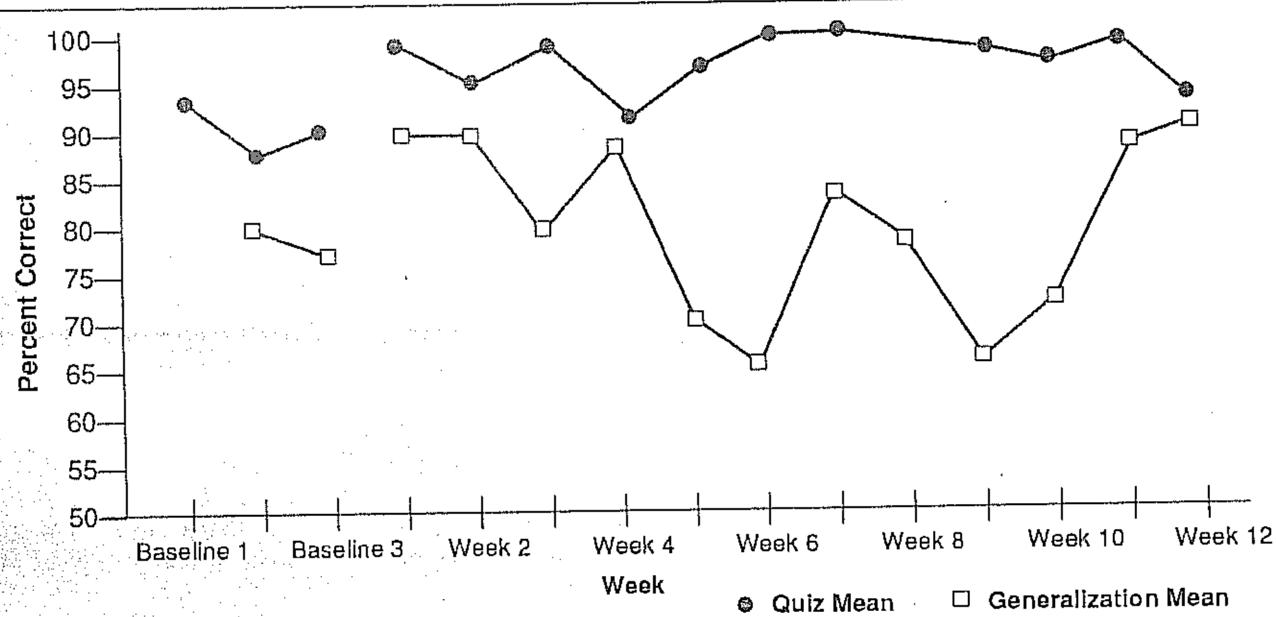
Two outcome measures were used: (1) student performance on Friday spelling tests, and (2) the percentage of previously practiced words spelled correctly on dictated sentences.

Results

The results are presented in Figure 1. Student performance on weekly quizzes averaged 90.2% correct during baseline conditions. The average for all weeks of CWPT was 97.1% correct.

Student performance on generalization probes during baseline conditions averaged 78.5%. This increased modestly to 79.7% during the 13 weeks of CWPT.

Figure 1. Mean Percent Correct, Spelling Quizzes and Generalization Probes.



Note: No spelling quiz administered in week 8, no generalization probe administered week 1.

Discussion

While we were clearly disappointed in the fact that CWPT did not improve student generalization of correct spelling compared to baseline condition. The lack of improvement should not obscure the fact that generalization was objectively good, nearly 80% correct usage on generalization tests.

III. Descriptive Analysis of Student Spelling Performance Within a Classwide Peer Tutoring System

Given the documented effects of CWPT on spelling test performance, and the known efficacy of it central instructional elements (frequent active responding and immediate feedback and error correction), we wondered how it is even possible for students not to master spelling words practiced. Clearly some students do not benefit from CWPT as much as others, and errors on the criterion measures (weekly spelling tests) continue to occurred. We, therefore, sought to examine possible factors which might contribute to student failure to benefit from CWPT. We hypothesized that students might fail to benefit from CWPT under any of the following conditions: (1) students have a lower rate of responding and therefore receive too little practice, (2) student response rates is adequate, but accuracy is very low, and/or (3) tutors do not use correct CWPT procedure (i.e., treatment fidelity is low) or participate in numerous competing behaviors.

Subjects

This investigation was conducted in a single classroom in a small city school district serving low SES students in western New York state. The study took place in a regular third grade classroom during scheduled spelling periods. Fourteen students (8M, 6F), ranging in age from 7 years 11 months to 10 years 6 months (mean = 8.9 years) served as subjects. Nine students were Hispanic, 2 were Black, 2 were Caucasian and 1 was Native American. Recent Stanford Achievement Test results indicated grade equivalents ranging from 1.6 to 3.7 in spelling (mean = 3.1).

The classroom teacher was highly experienced in the use of CWPT, having been part of a previous study five years earlier (Maheady & Harper, 1987), and had continued to use CWPT regularly for the intervening five years.

Procedure

CWPT was implemented, as described elsewhere, using 15 spelling words per day. Words were taken from the existing spelling text. Following four days of CWPT, a spelling test was administered on the words for that week. The teacher read each word orally, used it in a sentence, and pronounced it again.

Two types of data collection were undertaken. First, daily permanent products were collected and analyzed to determine: (a) the number of times each words was written or attempted (a practice measure), and (b) the percentage of correct responses. The second type of data collected was direct observation of tutoring dyads using an observational system developed by Kohler (1984). Core tutoring behaviors observed included: (a) verbal presentation of spelling words (e.g., "spell cat"), (b) verbal feedback on tutee's response (e.g., "correct" or "that's wrong"), (c) verbal administration of one or two points contingent upon tutee's response, and (d) verbal correction, if necessary. The tutee behavior observed was the verbal and written response to the "Spell items" task.

Other tutor behaviors observed included positive responses, neutral responses, negative responses, and off-task verbalizations. Two observers were used to obtain reliability data; interobservers were used to obtain reliability data; interobserver reliability averaged .91 for all trials.

Analysis of percent correct in spelling was done for all students. Observation were made only for those scoring lowest on the previous week's spelling test.

Results

To evaluate the first hypothesis, that students who failed to benefit from CWPT show low rates of responding, permanent products were examined for all students failing any spelling word on the Friday spelling test. Students' daily work products were analyzed retrospectively to determine their history with particular words spelled incorrectly on the subsequent exam. Results indicated that there were no significant differences in the rates of responding between students who failed to spell correctly one or more words and those who did not (mean = 4.46 and 4.57 respectively). Further, misspelled words were not written more or less often than those spelled correctly.

The second hypothesis, that when responding, lower achieving students would have lower accuracy rates, received partial support. For all students in a given week, percent correct varied between 45.6% and 94.5% with an overall average for all weeks of 76.45%. Among students who misspelled one or more words on the *Friday test*, the mean accuracy of responding during practice to those words was 68.12%. This varied between 0% and 100% correct during practice for individual students.

The third hypothesis was that students who made errors on the Friday test received incorrect CWPT procedure or engaged in numerous noncore behaviors. Examination of observational data indicated that key elements of the CWPT process were being omitted by participants. Among tutor behaviors, verbal presentation occurred in 99% of trials, verbal feedback to the tutee in 45.6% of trials and awarding of points in 70.7% of trials. The tutees verbal spelling of the word occurred in less than 1% of trials. The words were written each time, but not orally spelled. The omission of this step could account for the high rates of errors described above. Verbal correction occurred in only 9% of trials. Since the actual average error rate was 34%, many errors were going uncorrected.

Rates of all noncore behaviors were low. Positive behaviors occurred on 15% of trials, negative behaviors on 3.5% of trials, and off–task behaviors on 5.5% of trials.

Discussion

The present study involved a descriptive analysis of students' spelling performance under teacher–lead and classwide peer tutoring (CWPT) instructional conditions. An analysis of failures was conducted in an attempt to isolate practice conditions that may contribute to students' spelling failures. The results indicated that specific procedural omissions from the tutoring method (e.g., tutees' failure to spell words orally and tutors' failure to correct errors) may be responsible for subsequent spelling errors on weekly tests.

An analysis of students' interactions with subsequently failed words revealed that, most often, students practiced these words incorrectly during the week. While such an occurrence may be common under regular instructional conditions where immediate error correction by the teacher is virtually impossible, it was not anticipated under classwide peer tutoring. Subsequent direct observations of CWPT sessions among the lowest scoring students revealed substantial deviations from prescribed procedures. Verbal feedback from the tutor occurred on less than half of all dictated spelling trials. Points were awarded verbally less than three–fourths of the time. Most disconcerting, however, was the fact that tutees only verbally spelled and wrote dictated words 1% of the time. This is discouraging because it defeats the purpose of the oral spelling component of CWPT, that is; (a) to provide auditory reinforcement, and (b) to facilitate monitoring of spelling by the tutor. The findings suggest that on going monitoring of procedural implementation, particularly tutee responding and use of error corrections, is necessary. One would predict even higher student test performance given better error monitoring and correction.

Interpretation of the present findings must certainly be tempered by the use of a single classroom, rather small sample size, and the naturalistic design.

Summary and Conclusions

Collectively, findings from these three studies are significant in at least two regards. First, they add now information to the accumulating data base on a powerful, peer-mediated instructional intervention. It appears, for example, that classwide peer tutoring may be used by classroom teachers to: (1) increase student fluency on district-mandated instructional criteria (e.g., math computation rates), and/or (2) maintain acceptable levels of generalization among newly acquired spelling responses. In addition, CWPT procedures appear to lend themselves well to an analysis of treatment failure. this should permit both researchers and practitioners to identify procedural contributions to student failures within the tutoring system. In instances where procedural problems can be identified and controlled, student performance may be enhanced even further. ◆

References

Cook, N.L., Heron, T.E., & Heward, W.L. (1983). *Peer Tutoring: Implementing Classwide Programs in Primary Grades.* Columbus, OH: Special Press.

Delquadri, J., Greenwood, C.R., Stretton, K., & hall, R.V. (1983). The peer tutoring game: A classroom procedure for increasing opportunity to respond and spelling performance. *Education and Treatment of Children, 6,* 225–239.

Delquadri, J., Greenwood, C.R., Whorton, D., Carta, J.J., & Hall, R.V (1986). Classwide peer tutoring. *Exceptional Children, 52,* 535–542.

Greenwood, C.R, Delquadri, J. & Hall, R. (1984). Opportunity to respond and student academic performance in W.L. Heward, R.E. Herron, J. Trapp–Porter, & D.S. Hill (Eds.). *Focus on Behavior Analysis in Education* (pp. 58–88). Columbus OH: Charles, E. Merrill.

Greenwood, C.R., Delquadri, J.C., & Hall, R.V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology, 81,* 371–383.

Greenwood, C.R., Dinwiddie, G., Terry, B., Wade, L., Stanley, S., Thibadeau, S., & Delquadri, J.C. (1984). Teacher–versus peer–mediated instruction: An ecobehavioral analysis of achievement outcome. *Journal of Applied Behavior Analysis, 17,* 521–538.

Kohler, f. (1984). *Social Interaction Code for the Juniper Gardens Peer Tutoring Game.* Unpublished manuscript. University of Kansas, Juniper Gardens Project, Kansas City.

Maheady, L., & Harper, G.F. (1987). A classwide peer tutoring program to improve the spelling test performance of low income, thirds– and fourth–grade students. *Education and Treatment of Children, 10,* 120–133.

Maheady, L, Harper, G.F, & Sacca, M.K. (1988). The effects of classwide peer tutoring on the social studies test performance of secondary, mildly handicapped students in resource room programs. *Journal of Research and Development in Education, 2,* (3), 76–83.

Maheady, L., Sacca, M.K., & Harper, G.F. (187). Classwide student tutoring teams: The effects of peer mediated instruction on the academic performance of secondary mainstreamed students. *Journal of Special Education, 21,* (3), 107–121.

Maheady, L., Sacca, M.K., & Harper, G.F. (1988). Classwide peer tutoring with mildly handicapped high school students. *Exceptional Children, 55,* 52–59.

Designing Practice Activities*

by Douglas Carnine
University of Oregon

When reading "Information Processing and Elementary Mathematics" in the *Journal of Learning Disbilities*, by Pellegrino and Goldman (1987), I was struck by their call for research on "the form practice should take":

> Obviously, practice typically serves to strengthen knowlegde and it is the basic mechanism for explaining the acquisition of expertise . . . What is less obvious is the form that such practice should take; for example, item set size, distribution of practice, and errorless training, as well as amount of practice needed to produce rapid fact retrievial. (p. 32)

Pellegrino and Goldman (1987) and others such as Hasselbring, Goin, and Bransford (1988) relate practice to expertise via the concept of automaticity, in which students quickly call up basic information (e.g., the answer to 9 + 7, the sound that the letter *m* represents, of the name of a country on a map). Hasselbring et al. (1988) argue that all people have a limited capacity for information processing. When information can be accessed automatically, relatively little of this capacity is used, leaving ample capacity to focus on higher order skills. For example, a student counts on his fingers to answer 9 + 7 will be hampered in learning to carry, in a problem such as 249 + 137. An important goal of special education teachers, then, is to help their students develop automaticity of basic information.

This goal has been one line of research I have pursued over the past 15 years. In reading Pellegrino and Goldman's call for research, I realized that the procedures from those studies, appearing in various journals scattered over the years, had never been pulled together as guidelines for practitioners. Taken as a whole, the studies have numerous implications for "the form practice should take." The purpose of the present article is to review those studies, which evaluated ways to reduce errors during acquisition (thus reducing acquisition time), in order to increase accuracy after the practice and foster automaticity. The studies deal mostly with increasing accuracy and decreasing instructional time to mastery, both of which set the stage for automaticity. The studies sample a variety of topics—mathematics, reading, vocabulary, geography. The research also samples a variety of students—those with learning disabilities and those without—at different grade levels, elementary as well as secondary. (In all the studies, basic experimental protocol was followed, for example, students were randomly assigned to treatment.)

In drawing implications from such diverse studies, caution is appropriate. A procedure effective for designing practice for one topic with a particular type of student may not be effective for another topic with different students. Nevertheless, the procedures, as a group, provide a starting point for a coherent plan for designing practice procedures. The present review does not purport to be exhaustive; rather, the procedures are illustrative of those that have a research base. The studies have either been published, allowing the reader to locate additional references for procedures that are of particular interest, or have just recently been completed. The review is written for the practitioner who each day decides many times over "the form practice should take."

Avoid Memory Overload–Introduce Information Cumulatively

Conventional wisdom as well as research with college students suggests that the optimal number of new pieces of information to introduce to students is seven, plus or minus two (e.g., Hall, 1971). Seven is often too large for younger students who have learning disabilities. A preferred procedure is cumulative introduction, in which new pieces of information are introduced one at a time, after previously introduced information has been learned. Results from two studies support this recommendation.

Gleason, Carnine, and Vala (1988a) taught 47 elementary and junior high students with learning disabilities to identify the seven countries in Central America. In the cumulative introduction treatment, students discrminated two countries until six consecutive correct responses were made; then a third country was introduced. This cumulative introduction procedure was carried out until all seven countries had been introduced. In the rapid introduction treatment, after one country was introduced, the student identified it one time; and then the next country was introduced. This pattern was repeated seven times, resulting in all seven countries being introduced within a couple of minutes.

Both treatments continued until a student made 13 consecutive correct responses on a mix of all seven countries. The mean number of responses to criterion for the cumulative introduction treatment was 102; for the rapid introduction treatment, 258. The difference in minutes to criterion was comparable, 8.7 versus 22.7. Both differences were significant.

An experiment teaching six letter-sound correspondences to nonhandicapped preschoolers and first

*Published in the *Journal of Learning Disabilities*, vol. 10, pp. 603-607, 1989. Reprinted with the permission of the author and Pro-Ed.

tively, which meant that the time between the introduction of the first and last letters could have been two weeks. The order of introduction for the similar-separated group was *e, c, m, u, s, i* and the order for the similar-together group was *e, i, u, s, c, m*.

In Experiment 1, first graders in the similar-separated group made significantly more correct training responses than first graders in the similar-together group (51.7% versus 33.1%). Experiment 2 measured trials to criterion rather than number of correct responses following a fixed-trials presentation. Preschoolers in the similar-separated treatment reached criterion on *e* and *i* in significantly fewer trials, a mean of 178.0 versus 293.2 trials.

A related concern when designing practice activities is how to minimize the interference that results from the introduction of a highly similar letter. To investigate options for dealing with this problem, Carnine (1981) selected a parallelogram and its mirror image and labeled them *biff* and *diff*, to be analogous to *b* and *d*. Preschoolers were taught to identify examples of *biff, triangle* and *circle*, and then were randomly assigned to one of four treatments. The four treatments represented different methods of introducing *diff*, which was similar in appearance and name to the previously introduced *biff*. The terminal task for all treatments involved training to criterion on *biff, diff, circle*, and *triangle*. The four treatments were as follows:

1. No preteaching: Training consisted of just the terminal task. The other three methods involved preteaching.
2. Preteaching *diff* in the context of *circle* and *triangle:* Preschoolers labled examples of *cicrcle, triangle*, and *diff*, the new member. *Biff*, the previously introduced similar member, was excluded. In other words, preschoolers identified *diff* in the context of dissimilar members before discriminating *diff* from *biff*.
3. Preteaching *diff* and *biff*. The preschoolers labeled the examples as *biff* or *diff*, which required the preschoolers to make difficult discrminations based on both name and shape.
4. Preteaching *biff* and *not biff*. The preschoolers at first responded *biff* or *not biff*, allowing them to focus just on the shape discrimination because the discrimination involving similar names was postponed. The *diff* label was not introduced until the preschoolers reached criterion on *not biff*.

Preschoolers in the *biff* and *diff* pretraining treatment (3 above) required twice as many trials to reach criterion as preschoolers in any of the other three treatments, a mean of 30.5 versus 15.8 trials. The findings are consistent with the guideline of separating similar sound-symbol correspondences, in that the other three treatments all involved greater separation of the similar symbols than the *biff-diff* labeling treatment.

Make New Learning More Meaningful by Empahsizing Relationships

This procedure will be illustrated in three topic areas—reading, math, and geography. Each illustration is based on a different kind of relationship. For reading, it's teaching an algorithm to relate known letter-sound correspondences to words. For math, it's relating familiar facts to unknown facts, based on counting relationships. In geography, it's simply taking advantage of the physical arrangement of countries on a map.

Relate components to the whole. Carnine (1977) taught nonhandicapped preschoolers either letter-sound correspondences and a blending strategy for 18 words or rote identification of the 18 words. In the strategy treatment, students were shown how to use their knowledge of letter-sound correspondances to decode words. The memorization students didn't learn sounds or how sounds determine how a word is pronounced. They learned the 18 words through flash card drill. The strategy took 117 minutes to learn the letter-sound correspondences and master the 18 words. The memorization group took 132 minutes. Yet the strategy group was able to correctly identify significantly more new words, both phonically regular and irregular, than the memorization group.

Relate known information to unknown information. Thorton (1978) demonstrated the benefits of teaching addition facts as "an interrelated network" (Pellegrino & Goldman, 1987, p. 23). Carnine and Stein (1981) investigated a strategy for teaching easy facts (having addends of 1, 2, and 3) in whcih familiar facts (+ 1s) were used as an anchor for unfamiliar facts (+ 2s and + 3s). The strategy was based on the counting relationship between successive facts, indicated by the circled numbers that appear in the counting order.

$$6 + \textcircled{1} = 7$$

$$6 + \textcircled{2} = 8$$

$$6 + \textcircled{3} = 9$$

By requiring students to say the statements in order, teachers prompted students to recognize the counting relationship among the facts. For example, by knowing that 6 + 1 = 7, students can figure out the answer to 6 + 2: it's one more than 7. When an addend increases by one, the sum also increases by one.

In Study 1, preschoolers in both the strategy and memorization groups received training until they had mastered 21 facts, which meant that the amount of practice varied from student to student. Although differences in training time were not significant, differences in posttest scores favored the strategy group, 77% correct versus 33%. In Study 2, with first graders, a fixed amount of training was given to both groups.

Again, posttest differences significantly favored the strategy group.

The final study (Gleason, Carnine, & Valla, 1988b) used two different graphic organizers to teach elementary and secondary students with learning disabilities to identify seven Central American countries. In the spatially-related treatment, the countries were arranged as they appear on a map. In the random treatment, the countries were randomly distributed on the screen. By seeing the actual physical arrangement of the countries, students learned more than just to name the countries.

Students in the treatments required comparable times and numbers of responses to reach criterion. However, on the posttest, the students in the spatially related treatment had significantly higher scores.

Reduce Processing Demands

Presteach components of a strategy or algorithm. To sound out words, students need to know letter-sound correspondences. Although the correspondences could be introduced at the same time that students learn the sounding-out strategy, teaching the correspondences earlier would reduce the processing demands on students when they are trying to sound out words. The importance of preteaching component skills to the overall monitoring and orchestration of a complex operation is explained by Case (1975):

> In the early stages of skill acquisition, subjects have to monitor several external stimuli, and cooordinate a number of discrete responses. With overlearning, however, these responses become integrated into one unit. Since the skill is now much "simpler" from the subject's point of view, it consequently requires very little attention for execution. The result is (or should be) that a good deal of coordinating capacity is "left over," and that it can be used for integrating the newly consolidated basic skill with other basic skills. (p. 75)

Two studies investigated the effects of preteaching components, one with borrowing and one with multiplication. In the borrowing study (Kameenui & Carnine, 1986), second graders in the preteaching treatment learned where to borrow and how to rewrite a problem for borrowing, before actually working complete problems. For example, in the following task, students were instructed to cross out the box that shows where to borrow from:

```
 □□3□
 —□7□
```

Second, students practiced rewriting numbers such as 32 as 20 + 12. The no preteaching group learned the same algorithim for borrowing, but began working entire problems on the first day of instruction. Students in the preteaching group made significantly fewer errors during training, but had a score on the posttest comparable to that of the no preteaching treatment.

In the multiplication study (Carnine, 1980b), first graders who scored below average on a standardized readiness test received instruction in the spring of the year. The algorithm required the students to: (a) count by the number indicated by the first factor, and (b) count the number of times specified by the second factor. For example, in 5 x 3, the students would count by five, three times. (Students held up three fingers so they could keep track of how many times they had counted: 5, 10, 15.) In the preteaching treatment, students learned to count by various numbers, to translate problems into the appropriate action (e.g., raise the number of fingers indicated by the second factor), and how to coordinate touching a finger for each number that was counted. The no preteaching group learned these skills when working the 16 training problems. The preteaching group completed training in significantly less time (38 minutes versus 105 minutes), and scored significantly higher on eight new multiplication problems (61% versus 30%).

Teach easier information first. When learning a new type of information, such as the sound values for letters or how to chain together the steps in multiplication, processing demands increase. To keep students from experiencing demands that become so great as to impede learning, teachers should present easier-to-learn examples of the new type of information first.

Research on example difficulty has included both letters and words. Two experiments were conducted on the relative difficulty of six sound-symbol correspondences (Carnine, 1979)—three vowels (*e, i, u*) and three consonants (*c, s, m*). In the first experiment, eight preschoolers were trained to criterion on each of the letters in the context of three other letters; that is, six sets of letters were constructed, one for each of the six experimental letters.

The mean number of trials to criterion was 92.4 for *u*, 83.0 for *e*, 76.5 for *i*, 47.9 for *c*, 23.8 for *m*, and 11.6 for *s*. The vowels required significantly more trials than *m* and *s*. In the second experiment, 18 other preschoolers responded to the six experimental letters in a single set. Each letter appeared 100 times. The mean number of errors was 31.1 for *e*, 22.8 for *u*, 18.9 for *i*, 7.6 for *c*, 5.7 for *m*, and 1.7 for *s*. The vowels were significantly more difficult than the consonants, and *e* was significantly more difficult than *i*. The greater difficulty of the short vowels (*e, i, u*) in comparison to the consonants *c, s, m* is consistent with Coleman's (1970) findings. Note this caveat though: All the difficult letters shouldn't be bunched up at the end. The difficult letters, though not first in a sequence, will usually need to be distributed throughout the sequence.

In a study of word difficulty (Carnine & Carnine, 1978), children in a synthetic phonics program were tested on three types of untaught words: CVC (consonant-vowel-consonant) words with an initial stop sound, CVCC words with an initial continuous sound, and CCVC words. Students were presented with five examples of each word type (for a total of 15 words).

The mean percentage of correct responses was 71% for CVC words, 40% for CVCC words, and 22% for CCVC words. All differences were significant, suggesting that CVC words are easier to decode using a sounding-out strategy than CVCC and CCVC words, and that CVCC words are easier than CCVC words.

Require Quicker Responses

This is the only guideline that directly relates to automaticity. Hasselbring et al. (1988) call the procedure "controlling response times." This is the time allowed to retrieve and produce the answer to a fact. For math facts, they usually begin with a time of about 3 seconds and work down to a time of around 1.25 seconds. A response made after the allocated time is treated as an error. Obviously, this precise control of the time allowed for a response is best accomplished with a computer. Hasselbring et al. (1988) cautioned that, regardless of how the practice is delivered, by a person or by a computer, students should "know" the information before commencing this form of practice (e.g., a student should be able to produce the answer to 9 + 7 without having to count). The procedure of controlling response times has important implications for designing practice activities that foster automaticity.

Summary

When selecting or devising practice activities on basic information for students, teachers should consider these guidelines:

1. Prevent memory overload by introducing new information cumulatively.
2. Build retention with delayed review and discriminated practice.
3. Reduce interference effects by separating confusing elements from one another over time.
4. Emphasize relationships between components of an algorithm and the algorithm itself, between familiar and new information, and between items and their location, if relevant.
5. Reduce processing demands by introducing components before the algorithm or strategy itself and by introducing easier information first.
6. Require quicker responses to foster automaticity.

As noted in the introduction to this essay, these six guidelines have not been investigated across all topics or across all populations. Further research is needed. In the meantime, the guidelines serve as an initial response to Pellegrino and Goldman's (1987) question about the form practice should take for students with learning disabilities. The guidelines could be applied to a wide range of content for which automaticity is an ultimate goal: letter-sound correspondences, word reading, numeral identification, math facts, word meanings, geography facts, and so forth.

Authors Note. The preparation of this paper was supported in part by a grant from the Department of Education (G008730069). Suggestions by Geneva Blake and David Evans were quite helpful in revising the paper.

References

Carnine, D. W. (1976). Similar sound separation and cumulative introduction in learning letter-sound correspondences. *Journal of Educational Research, 69*, 368-372.

Carnine, D. W. (1977). Phonics versus look-say: Transfer to new words. *Reading Teacher, 30*(6), 636-640.

Carnine, D. W. (1979). Research on designing and implementing procedures for teaching sounds. In J. Button (Ed.), *Communication research in learning disabilities and mental retardation* (pp. 141-164). Baltimore: University Park Press.

Carnine, D. W. (1980a). Two letter discrimination sequences: High-confusion alternatives first versus low-confusion alternatives first. *Journal of Reading Behavior*, 12(1), 41-47.

Carnine, D. W. (1980b). Preteaching versus concurrent teaching of the component skills of a multiplication algorithm. *Journal of Research in Mathematics Education*, 11(5), 375-378.

Carnine, D. W. (1981). Reducing training problems associated with visually and auditorily similar correspondence. *Journal of Learning Disabilities, 14*, 276-279.

Carnine, D. W. & Stein, M. (1981). Organizational strategies and practice procedures for teaching basic facts. *Journal of Research in Mathematics Education*, 12(1), 65-69.

Carnine, L. & Carnine, D. W. (1978). Determining the relative decoding difficulty of three types of simple regular words. *Journal of Reading Behavior, 10*(4), 440-441.

Case, R. (1975). Gearing the demands of instruction to the developmental capacities of the learner. *Review of Educational Research, 45*, 59-87.

Coleman, E. B. (1970). Collecting a data base for reading technology. *Journal of Educational Psychology Monograph, 61*(4, Part 2).

Gleason, M., Carnine, D., & Valla, N. (1988a). *Cumulative introduction of information to students with learning disabilities*. Unpublished manuscript, University of Oregon, Eugene.

Gleason, M., Carnine, D., & Valla, N. (1988b). *Role of spatial organization in graphic organizers*. Unpublished manuscript, University of Oregon, Eugene.

Hall, J. F. (1971). *Verbal learning and retention*. Philadelphia: Lippincott.

Hasselbring, T., Goin, L., & Bransford, J. (1988). Developing math automaticity in learning handicapped children: The role of computerized drill and practice. *Focus on Exceptional Children, 20*(6), 1-15.

Johnson, G., Gersten, R., & Carnine, D. (1987). Effects of instructional design variables on vocabulary acquisition of LD students: A study of computer-assisted instruction. *Journal of Learning Disabilities, 10*, 206-213.

Kameenui, E. & Carnine, D. (1986). Preteaching versus concurrent teaching of the component skills of a subtraction algorithm to skill-deficient second graders: A component analysis of direct instruction. *The Exceptional Child*, 33(2), 103-115.

Kryzanowski, J. & Carnine, D. W. (1980). Effects of massed versus spaced formats in teaching sound/symbol correspondences to young children. *Journal of Reading Behavior*, 12, 225-230.

Pellegrino, J. W. & Goldman, S.R. (1987). Information processing and elementary mathematics. *Journal of Learning Disabilities, 20*, 23-32.

Thornton, A. (1978). Emphasizing thinking strategies in basic fact instruction. *Journal for Research in Mathematics Education*, 9, 214-227.

The Association for Direct Instruction Announces

The 16th Annual Eugene Direct Instruction Training and Information Conference

August 6–10, 1990

Eugene Hilton Hotel & Conference Center
Eugene, Oregon

Sessions

A Reading III-VI
A Arithmetic I & II
A Beginning Language
A Corrective Reading, Comprehension
A Effective Spelling Instruction
A Introduction to DI Techniques
A Overview of DI Research and Theory
A Solutions to Classroom Discipline Problems
A Instructional Techniques for Severely Handicapped Learners

B Reading I & II
B Reading III-VI
B Corrective Reading, Decoding
B Direct Instruction Math
B Supervision of Staff
B New Direct Instruction Programs
B Managing Serious Emotional Disturbances

C Reading Mastery & Literature
C Teaching Facts and Fact Systems
C Expressive Writing Instruction
C Options for At-risk and Special Needs Students
C DI Supplemental & Transitional Activities
C Summary of DI Research
C Teaching Study Skills
C Teach Your Child to Read in 100 Easy Lessons
C Advanced Teaching Techniques

Presenters:

Zig Engelmann, Wes Becker, Randy Sprick, Bob Dixon, Ann Arbogast, Jane Carter, Geoff Colvin, Gary Davis, Ann Glang, Mary Gleason, Phyllis Haddox, Tracey Hall, Gary Johnson, Marcy Stein, Linda Youngmayr

Space is limited! Register Now!

Conference Registration Information

Where-When: To be held August 6–10, 1990, at the Eugene Hilton Hotel and Conference Center, 66 East 6th Avenue, in downtown Eugene, Oregon.

How to Pre-Register: Please fill out the registration form. Enclose with check or Institutional purchase order for the proper fee. Send application to the Association for Direct Instruction. Pre-registration before July 7 guarantees space in preferred sessions. Any session with less than 20 participants may be cancelled. A confirmation receipt will be sent to all pre-registered participants. *THIS FORM COVERS CONFERENCE PRE-REGISTRATION ONLY. THIS DOES NOT CONSTITUTE PRE-REGISTRATION FOR COLLEGE CREDIT OR ROOM RESERVATION.*

Fees and Discounts: The conference registration fee is $175.00. Association members receive a 20% discount ($35.00 off). Groups of 5 to 9 participants receive a 10% discount. Groups of 10-19 receive a 20% discount. For groups of 20 or more, call for a quotation. Ask for Bryan Wickman at (503) 485-1293. The member and group discounts cannot be used together. Choose the discount that will benefit you the most. The fee does not include lodging or meals with the exception of the picnic, and coffee each morning. All training materials are included in the fee. New members are eligible for the 20% discount when membership application and appropriate fees accompany registration form.

Travel: The Association for Direct Instruction has selected Red Baron Travel as the Travel Agency for the Eugene Conference. On United Airlines flights, Red Baron can offer a 5% discount on the lowest applicable fare or 40% off coach fare, whichever is lower. Call Red Baron at 1-800-289-4222. You need to tell them you are attending the ADI Conference in order to get the discounted airfares.

Lodging: The special conference rate at the Eugene Hilton is $55.00 per day for a single, $62.00 double ($31.00 per person) plus tax. The Hilton has sold out for the past 6 years, so early reservations are recommended. You may contact the Hilton at (503) 342-2000. You need to tell them you are with the ADI group in order to receive the discounted room rates. There are a number of other hotels in the area. We will send a sheet of other lodging options to preregistrants. *DO NOT SEND ANY ROOM RESERVATION MONEY TO THE ASSOCIATION.*

College Credit: An optional 1, 2 or 3 hours of college credit through the University of Oregon Summer Session are available at an additional cost of $36.00 per quarter unit. The credit is listed as: Special Education 408G, Direct Instruction. Transcripts will be available in early October. Fee payment and registration will take place at the conference. *DO NOT SEND ANY COLLEGE CREDIT MONEY TO THE ASSOCIATION.*

Refunds and Cancellations: A 100% refund will be issued if a written request is postmarked by July 21. After that an 80% refund will be given. A written request must be received in our office before any refunds will be made.

Optional Events: Monday there will be a picnic at Skinners Butte Park to get acquainted. A meal for you and 1 guest is included in the registration fee. Wednesday from 4:00 to 5:00 there will be a special conference presentation. We will present the 1989 ADI Awards for Excellence in Education. Afterward, there will be a reception, providing an opportunity for conversation with trainers and other conference participants.

Conference Registration Form

Please fill out the registration form completely and mail to ADI.
Make checks payable (U.S. funds only) to **Association for Direct Instruction**
Because space is limited, early registration is recommended. Use an address where you will receive your mail up until the conference.

Name (as you would like it to appear on your name tag)________________ ______________________

Street ____

City ____________________ State ____________________ Zip ______________

Phone ______________________________

Agency Affiliation ______________________________ Position ______________

I would like to register for the following (list one "A", one "B", and one "C"):

"A" ______________________________

"B" ______________________________

"C" ______________________________

I am an Association for Direct Instruction member: Yes No
I will attend the picnic: Yes No
Have you attended the Eugene Conference previously?______ What year(s)? _______
Please return this form with your check or District Purchase Order to:

For office use only: Date_____ Fee_____ Check_____ PO#_______ By:____

ADI MATERIALS PRICE LIST

Theory of Instruction

Siegfried Engelmann & Douglas Carnine

Membership Price $20.00 List Price $25.00

Direct Instruction Reading

Douglas Carnine & Jerry Silbert

Membership Price $24.00 List Price $30.00

Direct Instruction Mathematics

Douglas Carnine, Marcy Stein & Jerry Silbert

Membership Price $24.00 List Price $30.00

Teach Your Child To Read in 100 Easy Lessons

Siegfried Engelmann, P. Haddox & E. Brunner

Membership Price $12.00 List Price $15.00

Generalized Compliance Training

Siegfried Engelmann & Geoff Colvin

Membership Price $16.00 List Price $20.00

Structuring Classrooms for Academic Success

S. Paine, J. Radicchi, L. Rosellini, L. Deutchman, C. Darch

Membership Price $8.00 List Price $10.00

Members of the Association for Direct Instruction may purchase copies of the materials listed above at the Membership price. Shipping charges are $1.50 per book for 1-5 books and $1.00 per book for orders of 6 or more. Orders are to be paid in U.S. Funds, in advance. Purchase orders are also accepted. Please allow 4 weeks for delivery.

ADI cannot provide copies for entire classes nor can we provide desk copies. All such requests must be made to the publisher of the specific book.

SEND YOUR CHECK OR PURCHASE ORDER TO:

Association for Direct Instruction
P.O. Box 10252
Eugene, OR. 97440
(503) 485-1293

"Why shood I lern too spel?"

—John Q. Student, 1990

Well, why should he? Presidents of companies can't. Famous athletes can't. Doctors are notoriously bad spellers. Rock artists don't bother. So, really, why should he?

For at least one important reason. Knowing how to spell facilitates clear, concise, and creative written expression. What if Charles Dickens hadn't been a very good speller? Would he have written "A Tale of Two Places" in which there were some "very bad times" instead of "A Tale of Two Cities" where it was "the best of times and the worst of times"?

When students don't know how to spell the words they want to use, they choose easier and less creative ways of expressing themselves.

SRA has a solution for breaking the spelling barrier. *Spelling Mastery. Spelling Mastery* is an innovative direct instruction program that emphasizes learning to spell by generalization rather than by the weekly memorization of word lists. *Spelling Mastery* teaches spelling strategies that transfer to writing because words are initially learned to a high level of mastery, and the program teaches transference to writing. This mastery approach works with all students, including remedial and special education students in grades 1-8.

For more information about *Spelling Mastery,* fill out the response form below and mail to: SRA, 155 N. Wacker Drive, Chicago, IL 60606.

SRA

☐ Yes! Please send me more information about *Spelling Mastery.*

☐ Please send me additional information about other Direct Instruction programs in:

☐ Reading ☐ Language ☐ Math

☐ Please have a sales representative contact me.

Name__________ Position__________

School__________

Address__________

County__________

Phone Number__________

Association for Direct Instruction
P.O. Box 10252
Eugene, OR. 97440

Non-Profit Organization
U.S. Postage Paid
Permit No. 122
Eugene, OR

Join the Association for Direct Instruction

Membership Options:

A) Regular Membership: $15.00 per year (includes one year of *DI NEWS* and a 20% discount on ADI sponsored items and events).

B) Student Membership: $7.00 per year (includes one year of *DI NEWS* and a 40% discount on ADI sponsored items and events).

C) Sustaining Membership: $30.00 or more per year (includes regular membership privileges and recognition of your support in the *DI NEWS*).

D) Institutional Membership: $50.00 per year (includes 5 subscriptions to the *DI NEWS* and membership privileges for 5 staff people).

E) *DI NEWS* Subscription only: $7.00 per year (outside North America and Hawaii $10.00 per year).

ADI sponsored products and events include books and other materials published or marketed by the Association for Direct Instruction. The ADI NEWS is published 4 times a year (Fall, Winter, Spring and Summer).

To join the Association, complete the bottom portion of this form and mail it, with your check or Purchase Order (in U.S. funds) to the Association for Direct Instruction.

Check one:
I wish to become an Association member. Please enroll me as a:

______ Regular Member ($15.00 annually)
______ Student Member ($7.00 annually)
______ Sustaining Member ($30.00 or more annually)
______ Institutional Membership ($50.00 annually)
______ I wish to subscribe to the DI NEWS only ($7.00 annually; $10.00 outside North America & Hawaii)

Name: ______________________________

Address: ______________________________

City, State, Zip: ______________________________

Phone: ______________________________

School District or Agency: ______________________________

Position: ______________________________