Alphabetics

Part II

Phonics
Instruction

PART II: PHONICS INSTRUCTION
Executive Summary

Introduction

Learning to read is a complex task for beginners. They must coordinate many cognitive processes to read accurately and fluently, including recognizing words, constructing the meanings of sentences and text, and retaining the information read in memory. An essential part of the process for beginners involves learning the alphabetic system, that is, letter-sound correspondences and spelling patterns, and learning how to apply this knowledge in their reading. Systematic phonics instruction is a way of teaching reading that stresses the acquisition of letter-sound correspondences and their use to read and spell words (Harris & Hodges, 1995). Phonics instruction is designed for beginners in the primary grades and for children having difficulty learning to read.

In teaching phonics explicitly and systematically, several different instructional approaches have been used. These include synthetic phonics, analytic phonics, embedded phonics, analogy phonics, onset-rime phonics, and phonics through spelling. Although all explicit, systematic phonics approaches use a planned, sequential introduction of a set of phonic elements along with teaching and practice of those elements, they differ across a number of other features. For example, the content covered ranges from a limited to an elaborate set of letter-sound correspondences and phonics generalizations. In addition, the application procedures taught to children vary. Synthetic phonics programs teach children to convert letters into sounds or phonemes and then blend the sounds to form recognizable words. Analytic phonics avoids having children pronounce sounds in isolation to figure out words. Rather children are taught to analyze letter-sound relations once the word is identified. Phonics-through-spelling programs teach children to transform sounds into letters to write words. Phonics in context approaches teach children to use sound-letter correspondences along with context cues to identify unfamiliar words they encounter in text. Analogy phonics programs teach children to use parts of written words they already know to identify new words. The distinctions between systematic phonics approaches are

not absolute, however, and some phonics programs combine two or more of these types of instruction. In addition, these approaches differ with respect to the extent that controlled vocabulary (decodable text) is used for practicing reading connected text. Although differences exist, the hallmark of systematic phonics programs is that they delineate a planned, sequential set of phonic elements and they teach these elements explicitly and systematically. The goal in all phonics programs is to enable learners to acquire sufficient knowledge and use of the alphabetic code so that they can make normal progress in learning to read and comprehend written language.

The purpose of this report is to examine the research evidence concerning systematic phonics instruction. The research literature was searched to identify experiments that compared the reading performance of children who had received systematic phonics instruction to the performance of children given nonsystematic phonics or no phonics instruction. The National Reading Panel (NRP) sought answers to the following questions:

- Does systematic phonics instruction help children learn to read more effectively than nonsystematic phonics instruction or instruction teaching no phonics?

- Are some types of phonics instruction more effective than others? Are some specific phonics programs more effective than others?

- Is phonics instruction more effective when students are taught individually, in small groups, or as whole classes?

- Is phonics instruction more effective when it is introduced in kindergarten or 1st grade to students not yet reading or in later grades after students have begun to read?

- Is phonics instruction beneficial for children who are having difficulty learning to read? Is it effective in preventing reading failure among children who are at risk for developing reading problems in the future? Is it effective in remediating reading

difficulties among children who have not made normal progress in learning to read?

• Does phonics instruction improve children's ability to read and comprehend text as well as their decoding and word-reading skills?

• Does phonics instruction have an impact on children's growth in spelling?

• Is phonics instruction effective with children at different socioeconomic (SES) levels?

• Does the type of instruction given to control groups as part of a study to evaluate phonics make a difference?

• If phonics instruction is found to be more effective than less-phonics or no-phonics instruction, were the experiments that showed these effects well designed or poorly designed?

Beginning reading programs that do not teach phonics explicitly and systematically may be of several types. In whole-language programs, the emphasis is upon meaning-based reading and writing activities. Phonics instruction is integrated into these activities but taught incidentally as teachers decide it is needed. Basal programs consist of a teacher's manual and a complete set of books and materials that guide the teaching of beginning reading. Some basal programs focus on whole-word or meaning-based activities with limited attention to letter-sound constituents of words and little or no instruction in how to blend letters to pronounce words. In sight word programs, children begin by building a reading vocabulary of 50 to 100 words, and then later they learn about the alphabetic system. These types of non-phonics programs were among those taught to children in the control groups of experiments examined by the NRP. Distinctions among the various types of non-phonics programs are not absolute. However, their defining characteristic is that they do not provide explicit, systematic phonics instruction.

Phonics programs have been used to teach young children to read as they progress through the primary grades and to remediate the reading difficulties of poor readers. The Panel analyzed studies that examined the effectiveness of phonics programs with three types of problem readers: children in kindergarten or 1st grade who were at risk for developing reading problems; older children of average or better intelligence who were not making normal progress in reading, referred to as

disabled readers; older children who were progressing poorly in reading and who varied in intelligence with at least some of them achieving poorly in other academic areas, referred to as low-achieving readers.

For children to learn to read, several capabilities must be developed. The focus of systematic phonics instruction is on helping children acquire knowledge of the alphabetic system and its use to decode new words, and to recognize familiar words accurately and automatically. Knowing how letters correspond to phonemes and larger subunits of words is essential for enabling beginning readers to sound out word segments and blend these parts to form recognizable words. Alphabetic knowledge is needed to figure out new words by analogy and to help beginners remember words they have read before. Knowing letter-sound relations also helps children to be more accurate in predicting words from context. In short, knowledge of the alphabetic system contributes greatly to children's ability to read words in isolation or connected text.

To study whether systematic phonics instruction improves children's ability to read words in various ways, different measures have been used. Decoding was tested by having children read regularly spelled words. To test whether children could read novel words, pseudowords (e.g., gan, bloff, trusk) were used. Sight vocabulary was examined through sets of leveled, miscellaneous words, not all of which were spelled regularly. In addition to word-reading, children's performance on measures of oral reading, text comprehension, and spelling was measured.

To provide solid evidence, experiments to test the contribution of systematic phonics instruction to reading acquisition must be well designed. Random assignment of students to treatment and control groups is a procedure that controls for other factors and allows researchers to conclude that the treatment itself was the cause of any growth in reading. However, sometimes the realities of schools and teachers make it impossible to randomly assign students, so researchers have to use quasi-experimental designs, assigning treatment and control conditions to already existing groups. Although researchers should administer pretests to determine whether the treatment and control groups differed prior to treatment and then remove any differences statistically when outcomes are analyzed, this is not always done. Also, larger sample sizes

provide more reliable findings, but access to many students is not always possible. In evaluating the evidence, the Panel attempted to rule out weak designs as the explanation for any positive effects that were produced by systematic phonics instruction.

Methodology

To evaluate the evidence, the NRP conducted a meta-analysis. The literature was searched electronically to locate potential studies. To qualify for the analysis, studies had to meet the following criteria:

1. Studies had to adopt an experimental or quasi-experimental design with a control group.

2. Studies had to appear in a refereed journal after 1970.

3. Studies had to provide data testing the hypothesis that systematic phonics instruction improves reading performance more than instruction providing unsystematic phonics or no phonics instruction. To be considered an instance of phonics instruction, the treatment had to teach children to identify or use symbol-sound correspondences systematically.

4. Studies had to measure reading as an outcome.

5. Studies had to report statistics permitting the calculation or estimation of effect sizes.

6. Studies were not those already included in the NRP's meta-analysis of phonemic awareness training studies.

From the potentially relevant list of references, 75 studies that appeared to meet the criteria were identified and located. These were carefully reviewed to determine their suitability for the meta-analysis. Studies of instructional interventions that might be found in schools were sought. Short-term laboratory studies and studies that taught only a limited set of processes were eliminated. Also eliminated were studies that simply compared different forms of phonics instruction but did not include a control group receiving reduced phonics or no phonics. Of the 75 studies screened, 38 were retained and 37 were eliminated from the final set used to calculate effect sizes.

The primary statistic used in the analysis of performance on outcome measures was effect size, indicating whether and by how much performance of the treatment group exceeded performance of the control group, with the difference expressed in standard deviation units. From the 38 studies entered into the database, 66 treatment-control group comparisons were derived.

Studies were coded for several characteristics that were included as moderators in the meta-analysis:

• Type of phonics program (synthetic programs emphasizing instruction in the sounding out and blending of words vs. programs teaching students to decode using larger subunits of words such as phonograms, as well as letters and sounds vs. miscellaneous programs),

• Specific phonics programs that were evaluated in at least three different studies (Direct Instruction; Lippincott; Orton Gillingham; Sing Spell Read and Write; Benchmark Word ID; New Primary Grades Reading System)

• Type of program taught to the control group (basal program, regular curriculum, whole language approach, whole word program, miscellaneous programs)

• Group assignment procedure (random assignment or nonequivalent groups)

• Number of participants (blocked into quartiles)

• Grade level (kindergarten, 1st grade, 2nd through 6th grades)

• Reading ability (normally developing, at risk, low achiever, reading disabled)

• Socioeconomic status (low, middle, varied, not given)

• Instructional delivery unit (class, small groups, 1:1 tutoring).

Children identified as being low achieving or at risk for reading failure were those tested and shown to have poor letter knowledge, poor phonemic awareness, or poor reading skills, or those in schools with low achievement, or those identified by teachers as needing special help in reading, or those who qualified for remedial programs in schools but the criteria for selection were not specified. Children classified as

reading disabled were those identified according to IQ-reading discrepancy criteria in standard use by researchers or those given tests to determine that the disability was reading-specific. In some cases, exclusionary criteria were applied as well (e.g., no neurological, behavioral, or emotional disorders).

Across the studies, the effects of phonics instruction on reading were most commonly assessed at the end of training. For programs lasting longer than one year, outcomes were measured at the end of each year in most cases. The primary outcome used in the meta-analysis was that assessed at the end of training or at the end of one year, whichever came first. Effect sizes were calculated on six types of outcome measures:

- Decoding regularly spelled real words
- Reading novel words in the form of pseudowords
- Reading miscellaneous words some of which were irregularly spelled
- Spelling words
- Comprehending text read silently or orally
- Reading text accurately aloud.

The mean effect size across these measures was calculated to yield a general literacy measure for each comparison. A statistical program was employed to calculate effect sizes and to test the influence of moderator variables on effect sizes. An effect size of $\underline{d} = 0.20$ is considered small; a moderate effect size is $\underline{d} = 0.50$; an effect size of $\underline{d} = 0.80$ or above is large.

Results and Conclusions

There were 38 studies from which 66 treatment-control group comparisons were derived. Although each comparison could contribute up to six effect sizes, one per outcome measure, few studies did. The majority (76%) of the effect sizes involved reading or spelling single words while 24% involved text reading. The imbalance favoring single words is not surprising given that the focus of phonics instruction is on improving children's ability to read and spell words. Moroever, many of the studies were conducted with beginning readers whose reading development at the time of the study was too limited to assess textual reading. Studies

limiting instructional attention to children with reading problems accounted for 65% of the comparisons, 38% involving poor readers considered at risk or low achieving, and 27% diagnosed as reading disabled (RD). Studies involving first graders were overrepresented in the database, accounting for 38% of the comparisons. Fewer kindergartners (12%) and children in 2nd through 6th grades (23%) were represented. Children in the RD group spanned several ages and grades, ranging from ages 6 to 13 and grades 2 through 6. Most of the studies (72%) were recent, conducted in the last 10 years.

Systematic phonics instruction typically involves explicitly teaching students a prespecified set of letter-sound relations and having students read text that provides practice using these relations to decode words. Instruction lacking an emphasis on phonics instruction does not teach letter-sound relations systematically and selects text for children according to other principles. The latter form of instruction includes whole word programs, whole language programs, and some basal reader programs.

The meta-analyses were conducted to answer several questions about the impact of systematic phonics instruction on growth in reading when compared to instruction that does not emphasize phonics. Findings provided strong evidence substantiating the impact of systematic phonics instruction on learning to read.

1. Does systematic phonics instruction help children learn to read more effectively than nonsystematic phonics instruction or instruction teaching no phonics?

Children's reading was measured at the end of training if it lasted less than a year or at the end of the first school year of instruction. The mean overall effect size produced by phonics instruction was moderate in size and statistically greater than zero, $\underline{d} = 0.44$. Findings provided solid support for the conclusion that systematic phonics instruction makes a bigger contribution to children's growth in reading than alternative programs providing unsystematic or no phonics instruction.

2. Are some types of phonics instruction more effective than others? Are some specific phonics programs more effective than others?

Three types of phonics programs were compared in the analysis: (1) synthetic phonics programs which emphasized teaching students to convert letters (graphemes) into sounds (phonemes) and then to blend the sounds to form recognizable words; (2) larger-unit phonics programs which emphasized the analysis and blending of larger subparts of words (i.e., onsets, rimes, phonograms, spelling patterns) as well as phonemes; (3) miscellaneous phonics programs that taught phonics systematically but did this in other ways not covered by the synthetic or larger-unit categories or were unclear about the nature of the approach. The analysis showed that effect sizes for the three categories of programs were all significantly greater than zero and did not differ statistically from each other. The effect size for synthetic programs was $\underline{d} = 0.45$, for larger-unit programs, $\underline{d} = 0.34$, and for miscellaneous programs, $\underline{d} = 0.27$. The conclusion supported by these findings is that various types of systematic phonics approaches are significantly more effective than non-phonics approaches in promoting substantial growth in reading.

There were seven programs that were examined in three or more treatment-control group comparisons in the database. Analysis of the effect sizes produced by these programs revealed that all were statistically greater than zero and none differed statistically from the others in magnitude. Effect sizes ranged from $\underline{d} = 0.23$ to 0.68. In most cases there were only three or four comparisons contributing effect sizes, so results may be unreliable. The conclusion drawn is that specific systematic phonics programs are all significantly more effective than non-phonics programs; however, they do not appear to differ significantly from each other in their effectiveness although more evidence is needed to verify the reliability of effect sizes for each program.

3. Is phonics taught more effectively when students are tutored individually or when they are taught in small groups or when they are taught as classes?

All three delivery systems proved to be effective ways of teaching phonics, with effect sizes of $\underline{d} = 0.57$ (tutoring), $\underline{d} = 0.43$ (small group), and $\underline{d} = 0.39$ (whole class). All effect sizes were statistically greater than zero, and no one differed significantly from the others. This supports the conclusion that systematic phonics instruction is effective when delivered through tutoring, through small groups, and through teaching classes of students.

4. Is phonics instruction more effective when it is introduced to students not yet reading, in kindergarten or 1st grade, than when it is introduced in grades above 1st after students have already begun to read?

Phonics instruction taught early proved much more effective than phonics instruction introduced after first grade. Mean effect sizes were kindergarten $\underline{d} = 0.56$; first grade $\underline{d} = 0.54$; 2nd through 6th grades $\underline{d} = 0.27$. The conclusion drawn is that phonics instruction produces the biggest impact on growth in reading when it begins in kindergarten or 1st grade before children have learned to read independently. These results indicate clearly that systematic phonics instruction in kindergarten and 1st grade is highly beneficial and that children at these developmental levels are quite capable of learning phonemic and phonics concepts. To be effective, systematic phonics instruction introduced in kindergarten must be appropriately designed for learners and must begin with foundational knowledge involving letters and phonemic awareness.

5. Is phonics instruction beneficial for children who are having difficulty learning to read? Is it effective in preventing reading failure among children who are at risk for developing reading problems in the future? Is it effective in remediating reading difficulties in children who have been diagnosed as reading disabled and children who are low-achieving readers?

Phonics instruction produced substantial reading growth among younger children at risk of developing future reading problems. Effect sizes were $\underline{d} = 0.58$ for kindergartners at risk and $\underline{d} = 0.74$ for 1st graders at risk. Phonics instruction also significantly improved the reading performance of disabled readers (i.e., children with average IQs but poor reading) for whom the effect size was $\underline{d} = 0.32$. These effect sizes were all statistically greater than zero. However, phonics instruction failed to exert a significant impact on the reading performance of low-achieving readers in 2nd through 6th grades (i.e., children with reading difficulties and possibly other cognitive difficulties explaining their low achievement). The effect size was $\underline{d} = 0.15$, which was not statistically greater than chance. Possible reasons might be that the phonics instruction provided to low-achieving readers was not sufficiently intense, or that their reading difficulties arose from sources not treated by phonics instruction such as poor comprehension, or there were too few cases (i.e., only eight treatment-control comparisons pulled from three studies) to yield reliable findings.

The conclusion drawn from these findings is that systematic phonics instruction is significantly more effective than non-phonics instruction in helping to prevent reading difficulties among at risk students and in helping to remediate reading difficulties in disabled readers. No conclusion is drawn in the case of low-achieving readers because it is unclear why systematic phonics instruction produced little growth in their reading and whether the finding is even reliable. Further research is needed to determine what constitutes adequate remedial instruction for low-achieving readers.

6. Does phonics instruction improve children's reading comprehension ability as well as their decoding and word-reading skills?

Systematic phonics instruction was most effective in improving children's ability to decode regularly spelled words ($\underline{d} = 0.67$) and pseudowords ($\underline{d} = 0.60$). This was expected because the central focus of systematic phonics programs is upon teaching children to apply the alphabetic system to read novel words. Systematic phonics programs also produced growth in the ability to read irregularly spelled words although the effect size was significantly lower, $\underline{d} = 0.40$. This is not surprising because a decoding strategy is less helpful for reading these words. However, alphabetic knowledge is useful for establishing connections in memory that help children read irregular words they have read before. This may explain the contribution of phonics.

Systematic phonics instruction produced significantly greater growth than non-phonics instruction in younger children's reading comprehension ability ($\underline{d} = 0.51$). However, the effects of systematic phonics instruction on text comprehension in readers above 1st grade were mixed. Although gains were significant for the subgroup of disabled readers ($\underline{d} = 0.32$), they were not significant for the older group in general ($\underline{d} = 0.12$).

The conclusion drawn is that growth in word-reading skills is strongly enhanced by systematic phonics instruction when compared to non-phonics instruction for kindergartners and 1st graders as well as for older struggling readers. Growth in reading comprehension is also boosted by systematic phonics instruction for younger students and reading disabled students. These findings should dispel the any belief that teaching phonics systematically to young children interferes with their ability to read and comprehend text. Quite the opposite is the case. Whether growth in reading comprehension is produced generally in students above 1st grade is less clear.

7. Does phonics instruction have an impact on children's growth in spelling?

Systematic phonics instruction produced much growth in spelling among the younger students, that is, kindergartners and 1st graders, $\underline{d} = 0.67$, but not among the older students (above 1st grade), whose effect size

of $\underline{d} = 0.09$ did not differ significantly from zero. One factor contributing to the difference is that younger children were given credit for using phonics-based knowledge to produce letter-sound spellings of words as well correct spellings whereas older children were not. Another factor may be that as children move up in the grades, remembering how to spell words requires knowledge of higher level regularities not covered in phonics programs. A third reason for the poor showing among older students may be that the majority were poor readers, known to have difficulty learning to spell.

The conclusion drawn is that systematic phonics instruction contributed more than non-phonics instruction in helping kindergartners and 1st graders apply their knowledge of the alphabetic system to spell words. However, it did not improve spelling in students above 1st grade.

8. Is phonics instruction effective with children at different SES levels?

Systematic phonics instruction helped children at all SES levels make significantly greater gains in reading than did non-phonics instruction. The effect size for low SES students was $\underline{d} = 0.66$ and for middle-class students was $\underline{d} = 0.44$. Both were statistically greater than zero and did not differ from each other. The conclusion drawn is that systematic phonics instruction is beneficial to students regardless of their SES.

9. Does the type of control group used to evaluate the effectiveness of phonics instruction make a difference?

The type of nonsystematic or non-phonics instruction given to control groups to evaluate the effectiveness of systematic phonics instruction varied across studies and included the following types: basal programs, regular curriculum, whole language approaches, whole word programs, and miscellaneous programs. The question of whether systematic phonics instruction produced better reading growth than each type of control group was answered affirmatively in each case. The effect sizes were all positive favoring systematic phonics, were all statistically greater than zero, and ranged from $\underline{d} = 0.31$ to 0.51. No single effect size differed from any of the others.

The conclusion supported by these findings is that the effectiveness of systematic phonics instruction found in the present meta-analysis did not depend on the type of instruction that students in the control groups received. Students taught phonics systematically outperformed students who were taught a variety of nonsystematic or non-phonics programs, including basal programs, whole language approaches, and whole-word programs.

10. Were studies reporting the largest effects of phonics instruction well designed or poorly designed experiments? That is, was random assignment used? Were the sample sizes sufficiently large? Might results be explained by differences between treatment and control groups that existed prior to the experiment rather than by differences produced by the experimental intervention?

The effects of systematic phonics instruction were not diminished when only the best designed experiments were singled out. The mean effect size for studies using random assignment to place students in treatment and control groups, $\underline{d} = 0.45$, was essentially the same as that for studies employing quasi-experimental designs, $\underline{d} = 0.43$, which used existing groups to compare phonics instruction and non-phonics instruction. The mean effect size for studies administering systematic phonics and non-phonics instruction to large samples of students did not differ from studies using the fewest students. For studies using between 80 and 320 students, $\underline{d} = 0.49$; for studies using between 20 and 31 students, $\underline{d} = 0.48$. There were some studies that did not use random assignment and either failed to address the issue of pre-existing differences between treatment and control groups or mentioned that a difference existed but did not adjust for differences in their analysis of results. The effect sizes changed very little when these comparisons were removed from the database, from $\underline{d} = 0.44$ to $\underline{d} = 0.46$.

The conclusion drawn is that the significant effects produced by systematic phonics instruction on children's growth in reading were evident in the most rigorously designed experiments. Significant effects did not arise primarily from the weakest studies.

11. Is enough known about systematic phonics instruction to make recommendations for classroom implementation? If so, what cautions should be kept in mind by teachers implementing phonics instruction?

Findings of the Panel regarding the effectiveness of systematic phonics instruction were derived from studies conducted in many classrooms with typical classroom teachers and typical American or English-speaking students from a variety of backgrounds and SES levels. Thus, the results of the analysis are indicative of what can be accomplished when systematic phonics programs are implemented in today's classrooms. Systematic phonics instruction has been used widely over a long period of time with positive results. A variety of phonics programs have proven effective with children of different ages, abilities, and socioeconomic backgrounds. These facts should persuade educators and the public that systematic phonics instruction is a valuable part of a successful classroom reading program. The Panel's findings summarized above serve to illuminate the conditions that make phonics instruction especially effective. However, caution is needed in giving a blanket endorsement to all kinds of phonics instruction.

It is important to recognize that the goals of phonics instruction are to provide children with some key knowledge and skills and to insure that they know how to apply this knowledge in their reading and writing. Phonics teaching is a means to an end. To be able to make use of letter-sound information, children need phonemic awareness. That is, they need to be able to blend sounds together to decode words, and they need to break spoken words into their constituent sounds to write words. Programs that focus too much on the teaching of letter-sounds relations and not enough on putting them to use are unlikely to be very effective. In implementing systematic phonics instruction, educators must keep the *end* in mind and insure that children understand the purpose of learning letter-sounds and are able to apply their skills in their daily reading and writing activities.

In addition to this general caution, several particular concerns should be taken into consideration to avoid misapplication of the findings. One concern relates to the commonly heard call for "intensive, systematic" phonics instruction. Usually the term "intensive" is not defined, so it is not clear how much teaching is required to be considered "intensive." Questions needing further answers are: How many months or years should a phonics program continue? If phonics has been taught systematically in kindergarten and 1st grade, should it continue to be emphasized in 2nd grade and beyond? How long should single instructional sessions last? How much ground should be covered in a program? That is, how many letter-sound relations should be taught and how many different ways of using these relations to read and write words should be practiced for the benefits of phonics to be maximum? These are among the many questions that remain for future research.

Secondly, the role of the teacher needs to be better understood. Some of the phonics programs showing large effect sizes are scripted in such a way that teacher judgment is largely eliminated. Although scripts may standardize instruction, they may reduce teachers' interest in the teaching process or their motivation to teach phonics. Thus, one concern is how to maintain consistency of instruction and at the same time encourage unique contributions from teachers. Another concern involves what teachers need to know. Some phonics programs require a sophisticated understanding of spelling, structural linguistics, and word etymology. Teachers who are handed the programs but are not provided with sufficient inservice training to use these programs effectively may become frustrated. In view of the evidence showing the effectiveness of systematic phonics instruction, it is important to ensure that the issue of how best to prepare teachers to carry out this teaching effectively and creatively is given high priority. Knowing that all phonics programs are not the same brings with it the implication that teachers must themselves be educated about how to evaluate different programs, to determine which are based on strong evidence and how they can most effectively use these programs in their own classrooms.

As with any instructional program, there is always the question: "Does one size fit all?" Teachers may be expected to use a particular phonics program with their class, yet it quickly becomes apparent that the program suits some students better than others. In the early

grades, children are known to vary greatly in the skills they bring to school. There will be some children who already know most letter-sound correspondences, some children who can even decode words, and others who have little or no letter knowledge. Should teachers proceed through the program and ignore these students? Or should they assess their students' needs and select the types and amounts of phonics suited to those needs? Although the latter is clearly preferable, this requires phonics programs that provide guidance in how to place students into flexible instructional groups and how to pace instruction. However, it is common for many phonics programs to present a fixed sequence of lessons scheduled from the beginning to the end of the school year.

Finally, it is important to emphasize that systematic phonics instruction should be integrated with other reading instruction to create a balanced reading program. Phonics instruction is never a total reading program. In 1st grade, teachers can provide controlled vocabulary texts that allow students to practice decoding, and they can also read quality literature to students to build a sense of story and to develop vocabulary and comprehension. Phonics should not become the dominant component in a reading program, neither in the amount of time devoted to it nor in the significance attached. It is important to evaluate children's reading competence in many ways, not only by their phonics skills but also by their interest in books and their ability to understand information that is read to them. By emphasizing all of the processes that contribute to growth in reading, teachers will have the best chance of making every child a reader.

Directions for Further Research

Although phonics instruction has been the subject of a great deal of study, there are important topics that have received little or no research attention, and there are other topics that, although previously studied, require further research to refine our understanding.

Three important but neglected questions are prime candidates for research: What are the "active ingredients" in effective systematic phonics programs? Is phonics instruction improved when motivational factors are taken into account—not only learners' but also teachers' motivation to teach? How does the use of decodable text as early reading material contribute to the effectiveness of phonics programs?

1. Active Ingredients

Systematic phonics programs vary in many respects. It is important to determine whether some properties are essential and others are not. Because instructional time during the school day is limited, teachers and publishers of beginning reading programs need to know which ingredients of phonics programs yield the most benefit.

2. Motivation

Phonics instruction has often been portrayed as involving "dull drill" and "meaningless worksheets." Few if any studies have investigated the contribution of motivation to the effectiveness of phonics programs, not only the learner's motivation to learn but also the teacher's motivation to teach. The lack of attention to motivational factors by researchers in the design of phonics programs is potentially very serious because debates about reading instruction often boil down to concerns about the "relevance" and "interest value" of *how* something is being taught, rather than the specific content of *what* is being taught. Future research on phonics instruction should investigate how best to motivate children in classrooms to learn the letter-sound associations and to apply that knowledge to reading and writing. It should also be designed to determine which approaches teachers prefer to use and are most likely to use effectively in their classroom instruction.

3. Decodable Text

Some systematic phonics programs are designed so that children are taught letter-sound correspondences and then provided with little books written carefully to contain the letter-sound relations that were taught. Some programs begin with a very limited set and expand these gradually. The intent of providing books that match children's letter-sound knowledge is to enable them to experience success in decoding words that follow the patterns they know. The stories in such books often involve pigs doing jigs and cats in hats. Systematic phonics programs vary in the percentage of decodable words in 1st-grade stories and in the

percentage of sight words introduced holistically to make a good story. Surprisingly, very little research has attempted to determine the contribution of decodable books to the effectiveness of phonics programs.

There are other important topics to be addressed in future research as well. These include the following:

•   Should systematic phonics instruction continue beyond 2nd grade? If so, what are the goals of more advanced forms of phonics instruction and does this instruction contribute to growth in reading?

•   Are there ways to improve the effectiveness of systematic phonics instruction for poor readers above 1st grade? Does this instruction need to take account of any maladaptive reading habits the students have acquired or any sources impeding the incorporation of alphabetic knowledge and decoding strategies into their reading? Does this instruction need to take account of the type of reading instruction they experienced in earlier years? Does decoding instruction need to be combined with comprehension instruction?

Report

Introduction

Learning to read is a complex task for beginners. They must coordinate many cognitive processes to read accurately and fluently. Readers must be able to apply their alphabetic knowledge to decode unfamiliar words and to remember how to read words they have read before. When reading connected text, they must construct sentence meanings and retain them in memory as they move on to new sentences. At the same time, they must monitor their word recognition to make sure that the words activated in their minds fit with the meaning of the context. In addition, they must link new information to what they have already read, as well as to their background knowledge, and use this to anticipate forthcoming information. When one stops to take stock of all the processes that readers perform when they read and comprehend text, one is reminded how amazing the act of reading is and how much there is for beginners to learn.

In teaching phonics explicitly and systematically, several different instructional approaches have been used. These include synthetic phonics, analytic phonics, embedded phonics, analogy phonics, onset-rime phonics, and phonics through spelling. Although these explicit and systematic phonics approaches all use a planned, sequential introduction of a set of phonic elements with teaching and practice of those elements, they differ across a number of other features. For example, the content covered ranges from a limited to an elaborate set of letter-sound correspondences and phonic generalizations. The application procedures taught to children vary. Synthetic phonics programs teach children to convert letters into sounds or phonemes and then blend the sounds to form recognizable words. Analytic phonics avoids having children pronounce sounds in isolation to figure out words. Rather, children are taught to analyze letter-sound relations once the word is identified. Phonics-through-spelling programs teach children to transform sounds into letters to write words. Phonics in context approaches teach children to use sound-letter correspondences along with context cues to identify unfamiliar words they encounter in text. Analogy phonics programs teach children to use parts

of written words they already know to identify new words. The distinctions between systematic phonics approaches are not absolute, however, and some phonics programs combine two or more of these types of instruction. In addition, these approaches differ with respect to the extent that controlled vocabulary (decodable text) is used for practicing reading connected text. Although these differences exist, the hallmark of systematic phonics programs is that they delineate a planned, sequential set of phonic elements, and they teach these elements, explicitly and systematically. The goal is to enable learners to acquire sufficient knowledge and use of the alphabetic code so that they can make normal progress in learning to read and comprehend written language.

A key feature that distinguishes systematic phonics instruction from nonsystematic phonics is in the identification of a full array of letter-sound correspondences to be taught. The array includes not only the major correspondences between consonant letters and sounds but also short and long vowel letters and sounds, and vowel and consonant digraphs (e.g., oi, ea, ou, sh, ch, th). Also, it may include blends of letter-sounds that recur as subunits in many words, such as initial blends (e.g., st, sm, bl, pr), and final stems (e.g., -ack, -end, -ill, -op). Learning vowel and digraph spelling patterns is harder for children; therefore, special attention is devoted to learning these relations. It is not sufficient just to teach the alphabetic system. Children need practice in applying this knowledge in reading and writing activities. Programs provide practice in various ways. Phonics programs may teach children decoding strategies that involve sounding out and blending individual letters and digraphs, or pronouncing and blending larger subunits such as initial blends and final stems of words. Programs may provide children with text whose words can be decoded using the letter-sound relations already taught. Programs may have children write their own text using the letter-sounds taught and then have children read their own and others' stories.

The purpose of literacy instruction in schools is to help children master the many challenges of written language. While teachers use a variety of activities to accomplish this purpose, one central approach is to teach the alphabetic code that represents oral language in writing. Children need to understand how letters, called graphemes, stand for the smallest sounds, called phonemes, in spoken words. Systematic phonics instruction teaches beginning readers the alphabetic code consisting of a large set of correspondences between graphemes and phonemes and perhaps larger sub-units of words and how to use this knowledge to read words. In some phonics programs, beginners are taught a routine for transforming spellings into blends of phonemes that are recognized as words. Learning about letter-sound associations helps beginners break the code in learning to read. However, the English writing system has other higher level, word-based regularities as well, so, although phonics instruction contributes, it is not the complete solution to word identification that it is in other written languages that are more fully phonemic (e.g., Spanish).

Over the years educators have disagreed about how beginning reading should be taught. Some have advocated starting with a systematic phonics approach while others have argued for a whole word approach or a whole language approach. Disagreement has centered on whether teaching should begin with systematic explicit instruction in symbol-sound correspondences, whether it should begin with whole words, or whether initial instruction should be meaning-centered with correspondences taught incidentally in context as needed. Most recently the pendulum has swung toward providing children with more explicit phonics instruction. Educators advocating this shift have claimed that there is substantial research showing that approaches with an emphasis on phonics instruction are more effective than approaches that do not emphasize the teaching of phonics.

The purpose of this report was to examine the research evidence concerning phonics instruction. The Panel sought answers to the following questions:

- Does systematic phonics instruction help children learn to read more effectively than unsystematic phonics instruction or instruction teaching no phonics?

- Are some types of phonics instruction more effective than others? Are some specific phonics programs more effective than others?

- Is phonics instruction more effective when it is introduced to students not yet reading, in kindergarten or 1st grade, than when it is introduced in grades above 1st after students have already begun to read?

- Is phonics instruction beneficial for children who are having difficulty learning to read? Is it effective in preventing reading failure among children who are at risk for developing reading problems in the future? Is it effective in remediating reading difficulties among children who have not made normal progress in learning to read?

- Is phonics taught more effectively when students are tutored individually, or when they are taught in small groups, or when they are taught as classes?

- Does phonics instruction improve children's ability to read connected text as well as their decoding and word reading skills?

- Does phonics instruction have an impact on children's growth in spelling?

- Is phonics instruction effective with children at different socioeconomic levels?

- Does the type of instruction given to control groups and used to evaluate the effectiveness of phonics instruction make a difference? That is, is systematic phonics more effective than forms of instruction that do not emphasize phonics, such as the whole word approach or meaning-centered approaches?

- If phonics instruction is found to be more effective than less-phonics or no-phonics instruction, were the experiments showing these effects well designed or poorly designed?

To evaluate the evidence, a meta-analysis was conducted. The Panel searched the literature to locate experimental studies published after 1970 that administered systematic phonics instruction to one group of children and administered another type of instruction that involved unsystematic phonics or no phonics to a control group. Also the studies had to examine phonics programs of the sort used in schools rather than single-process-focused laboratory procedures. The studies had to measure reading as an

outcome of instruction. In addition, studies were excluded if they were in the Panel's other database used to conduct a meta-analysis examining effects of phonemic awareness instruction on reading. A total of 38 studies meeting the NRP research criteria was found. The studies were coded for various characteristics of students, instruction, and experimental design. A meta-analysis was conducted to examine the size of effects that resulted when the performance of students receiving systematic phonics instruction was compared to that of students receiving another form of instruction that did not focus on phonics. The outcomes measured following instruction included children's ability to read words and pseudowords, to read and comprehend text, and also to spell words.

Background and Rationale for the Meta-Analysis

Historical Overview

The question of whether instruction that includes an initial emphasis on systematic phonics is more effective than other forms of instruction in teaching children to read has been addressed many times in the literature. The particular issues underlying interest in this question have shifted over the years, but the topic has remained controversial, and this has spawned a number of reviews of research.

In the 1960s, the Office of Education funded the Cooperative Research Program in First Grade Reading (Bond & Dykstra, 1967, 1998) and Project Literacy (Levin & Williams, 1970). The First Grade studies involved a wide-ranging research project, consisting of 29 separate studies in different sites, all aimed at determining the "best" approach to teaching beginning reading. In contrast, Project Literacy attempted to identify the basic psychological and linguistic processes involved in learning to read and did not focus directly on the pedagogy of reading. At the same time, the Carnegie Foundation funded Jeanne Chall's (1967) comprehensive review of beginning reading instruction, Learning to Read: The Great Debate. That review, like the present report, was intended to analyze the results of previous research.

Concern about beginning reading instruction was not confined just to the educational community but was very much in public discourse. Flesch (1955) had authored a best selling book Why Johnny Can't Read in

which he argued that children were being abused by the then-current whole word methodology. Flesh asserted that if children were taught only the 44 letter-sound correspondences, they would be able to read any word they encountered, and there would be no reading problems. Spurred on partially by Flesch and partially by advances in linguistics, new phonics programs were developed and began achieving wider usage in reading instruction (Aukerman, 1981; Popp, 1975).

Chall's (1967) review examined both the underlying theory and the classroom realities of these new phonics programs. But the core of her study was a comprehensive analysis of the research up to the mid-1960s, including the then-unpublished First Grade Studies. Chall's basic conclusion continues to be cited to this day, her finding that early and systematic instruction in phonics seems to lead to better achievement in reading than later and less systematic phonics instruction.

It is important to note that Chall, in the 1967 edition of her review, did not recommend any particular type of phonics instruction. Common forms of phonics instruction in the 1960s included synthetic instruction, analytic instruction, and linguistic readers (Aukerman, 1981). All of these challenged the sight word approach of the day. However, in the 1983 edition of her review, Chall did suggest that synthetic phonics instruction held a slight edge over analytic phonics instruction. Even in this, her recommendation was temperate.

Chall's (1967) basic finding has been reaffirmed in nearly every research review conducted since then (e.g., Adams, 1990; Anderson et al., 1985; Balmuth, 1982). Also, one of the coordinators of the First Grade Studies (Dykstra, 1968) published an analysis in which he concluded that the results of that project supported Chall's basic finding (Adams, 1990). Nevertheless, the controversy has persisted over this issue (Grundin, 1994; Taylor, 1998; Weaver, 1998). Part of the reason that the debate has continued is that phonics instruction has become entangled with politics and ideology (Goodman, 1993; McKenna, Stahl, & Reinking, 1994; Stahl, 1999). Another reason has been philosophical disagreements about how children learn to read and confusions about the implications of these varied points of view.

Phonics and No-Phonics Instruction

At the time of Chall's (1967) original review, the contrast between phonics and the alternative "look-say" methods was considerable. In the look-say approach, children were taught to read words as wholes much like Chinese logographs, and they practiced reading words until they had acquired perhaps 50 to 100 words in their sight vocabularies. Only after this accomplishment, which occurred toward the end of 1st grade, did phonics instruction begin. This was truly non-phonics instruction because discussion of letter-sound relations was delayed for a considerable length of time. The look-say approach contrasted with a variety of phonics programs. These included synthetic phonics programs which taught children to sound out and blend words, linguistic programs which taught decoding through patterned words and phonetically controlled texts, and analytic phonics programs which taught children to analyze letter-sound relations in previously learned words so as to avoid pronouncing sounds in isolation (Aukerman, 1971, 1984).

In the present day, whole language approaches have replaced the whole word method as the alternative to systematic phonics programs. The shift has involved a change from very little letter-sound instruction in 1st grade to a modicum of letter-sounds taught unsystematically. In contrast to the whole word method, whole language teachers are not told to wait until a certain point before teaching children about letter-sound relationships. Whereas in the 1960s, it would have been easy to find a 1st grade reading program without any phonics instruction, in the 1980s and 1990s this would be rare. Baumann, Hoffman, Moon, and Duffy-Hester (1998), in a national survey of 1,207 elementary school teachers, found that 63% believed that phonics should be taught directly and that 89% believed that skills instruction should be combined with literature and language-rich activities. Fisher, Lapp, and Flood (1999), in a survey of 118 California teachers, found that 64% of the K through 2 teachers integrated phonics instruction into their lessons (with some extra isolated phonics), and the remainder taught phonics as a separate part of word study.

Whole language teachers typically provide some instruction in phonics, usually as part of invented spelling activities or through the use of graphophonemic prompts during reading (Routman, 1996). However,

their approach is to teach it unsystematically and incidentally in context as the need arises. The whole language approach regards letter-sound correspondences, referred to as graphophonemics, as just one of three cueing systems (the others being semantic/meaning cues and syntactic/language cues) that are used to read and write text. Whole language teachers believe that phonics instruction should be integrated into meaningful reading, writing, listening, and speaking activities and taught incidentally when they perceive it is needed. As children attempt to use written language for communication, they will discover naturally that they need to know about letter-sound relationships and how letters function in reading and writing. When this need becomes evident, teachers are expected to respond by providing the instruction.

Although some phonics is included in whole language instruction, important differences have been observed distinguishing this approach from systematic phonics approaches. In several vignettes portraying phonics instruction in whole language contexts (Dahl, Sharer, Lawson, & Grogran, 1999; Freppon & Dahl, 1991; Freppon & Headings, 1996; Mills, O'Keefe, & Stephens, 1992), few if any instances of vowel instruction were found (Stahl, Duffy-Hester, & Stahl, 1998). This contrasts with systematic phonics programs where the teaching of vowels is central and is considered essential for enabling children to decode (Shankweiler & Liberman, 1972).

Another practice that is found in some systematic phonics programs but is not found in whole language programs is that of teaching children to say the sounds of letters and blend them to decode unfamiliar words. Programs that teach this procedure are referred to as synthetic phonics programs. Systematic phonics programs also commonly teach children an extensive, pre-specified set of letter-sound correspondences or phonograms while whole language programs teach a more limited set, in context, as needed. Systematic phonics programs teach phonics explicitly by delineating a planned, sequential set of phonic elements and teaching these elements explicitly and systematically; some systematic phonics programs also use controlled vocabulary (decodable text) to provide practice with these elements. Whole language programs do not

prespecify the relations to be taught. It is presumed that exposing children to letter-sound relations as they read text will foster incidental learning of the relations they need to develop as readers.

The meta-analysis was conducted to compare the effectiveness of systematic phonics instruction to other forms of instruction lacking an emphasis on phonics. Included in the database were several studies that provided whole language instruction to control groups and studies teaching whole word programs to control groups. In fact, two studies in the database were conducted for the purpose of evaluating the effects of whole language programs, not phonics programs. In these studies, phonics was the form of instruction given to control groups (Klesius et al., 1991; Freppon, 1991).

Not only whole language and whole word instruction but also other forms of control-group instruction were present in the database. Several control groups received some type of basal instruction, usually a program prescribed by the school or district. Basal programs consist of a whole package of books and supplementary materials that are used to teach reading. Teachers work from a thick manual that details daily lesson plans based on a scope and sequence of the reading skills to be taught. Students are given workbooks to practice on skills. Tests are used to place students in the proper levels of the program and to assess mastery of skills (Aukerman, 1981). Basal reading programs do vary, but one can assume that basal readers of the same era are roughly similar in their characteristics. The basal programs given to control groups provided only limited or no systematic phonics instruction.

A few studies utilized as their baseline control the performance of comparable classes of students enrolled in the same schools the year prior to the treatment (Snider, 1990; Vickery et al., 1987). In one case, a basal program was used. In the other case, the type of program was not specified. Campbell and Stanley (1966) suggest that this design contains certain threats to external validity, especially the differential history of the two groups.

Some studies in the database included more than one control group. The Panel selected for the meta-analysis the group receiving the least phonics instruction.

The issue of the control group is crucial. A meta-analysis compares a treatment to what is supposedly a constant. However, in reality, the size of the effect is a result of what goes on in both the treatment and the control groups. A treatment can be very effective but yield only a small effect size if instruction in the control group is also effective. On the other hand, if the control group's instruction is particularly ineffective, by design or by accident, then the effect size is inflated. One must consider the nature of the control group in order to interpret an effect size. The question addressed in the meta-analysis was whether phonics instruction produced greater growth in reading than each of the various types of instruction given to control groups.

Types of Phonics Instruction

The hallmarks of systematic phonics programs are that children receive explicit, systematic instruction in a set of prespecified associations between letters and sounds, and they are taught how to use them to read, typically in texts containing controlled vocabulary. However, phonics programs vary considerably in exactly what children are taught and how they are taught (Adams, 1990; Aukerman, 1981). Approaches to phonics instruction may differ in several important ways including the following:

1. How many letter-sound relations are taught, how they are sequenced, whether phonics generalizations are taught as well (e.g., "When there are two vowels side by side, the long sound of the first one is heard and the second is usually silent."), whether special marks are added to letters to indicate their sounds, for example, curved or straight lines above vowels to mark them as short or long

2. The size of the unit taught (i.e., graphemes and phonemes, or larger word segments called phonograms, for example, -ing, or -ack which represent the rimes in many single-syllable words)

3. Whether the sounds associated with letters are pronounced in isolation (synthetic phonics) or only in the context of words (analytic phonics)

4. The amount and type of phonemic awareness that is taught, for example, blending or segmenting sounds orally in words

5. Whether instruction is sequenced according to a hierarchical view of learning with the steps regarded as a series of prerequisites (i.e., letters, then letter-sound relations, then words, then sentences) or whether multiple skills are learned together

6. The pace of instruction

7. The word reading operations that children are taught, for example, sounding out and blending letters, or using larger letter subunits to read words by analogy to known words

8. The involvement of spelling instruction

9. Whether learning activities include extensive oral drill-and-practice, reciting phonics rules, or filling out worksheets

10. The type of vocabulary control provided in text (e.g., is the vocabulary limited mainly to words containing familiar letter-sound associations or are sight words introduced to help create a meaningful story?)

11. Whether phonics instruction is embedded in or segregated from the literacy curriculum

12. The teaching approach, whether it involves direct instruction in which the teacher takes an active role and students passively respond, or whether a "constructivist" approach is used in which the children learn how the letter-sound system works through problemsolving

13. How interesting and motivating the instructional activities are for teachers and for students.

Systematic phonics programs included in the Panel's database varied in many of these ways; so, it should not be assumed that the programs taught phonics uniformly. One purpose of the meta-analysis was to examine whether different properties of phonics programs influenced how effective they were in teaching children to read. However, this purpose was thwarted by the fact that most studies did not describe the phonics instruction in sufficient detail to permit coding the properties listed above. As a result, the Panel selected only one property for coding: whether programs emphasized a synthetic approach in teaching children to read words or whether the emphasis was on larger subunits of words.

A majority of the programs in the database used a synthetic approach to teach phonics. This instruction typically begins by teaching children relations between individual letters and pairs of letters called digraphs (e.g., TH, AI, CH, OI) and all 44 sounds or phonemes of the language. These correspondences are introduced systematically and sequentially. Children are taught to decode unfamiliar words by sounding out the letters and blending them to pronounce a recognizable word.

However, the synthetic strategy presents two difficulties for children. One is that blending words containing stop consonants requires deleting "extra" (schwa vowel) sounds produced when letters are pronounced separately, for example, blending "tuh-a-puh" requires deleting the "uh" sounds to produce the blend "tap." The second problem is that when the sounds to be blended exceed two or three, it becomes harder to remember and manage the ordering of all those sounds, for example, blending "s-tuh-r-ea-m" to say "stream."

Phonics programs have been developed to address these difficulties. One approach used has been to teach students to read larger subunits of words as well as phonemes. For example, children learn to recognize ST, AP, EAM, as blends so that there are not so many separate parts of words to sound out and remember in blending them. The larger units taught might include onsets (i.e., the consonants that precede the vowel such as "st" in stop) and rimes (i.e., the vowel and following consonants such as "op" in stop), also called phonograms, and spelling patterns characterizing the common parts of word families (e.g., -ack as in pack and stack, -oat as in goat and float). Teaching children to analyze and pronounce parts of words provides the basis for teaching them the strategy of reading new words by analogy to known words (e.g., reading stump by analogy to jump). In the database, these studies are distinguished and classified as teaching children to analyze and blend words by using larger phonological units.

The database included 43 treatment-control comparisons that taught synthetic phonics to the treatment groups, 11 studies that used phonics treatments emphasizing larger subunits for blending words, two comparisons that combined both types of

programs, and ten comparisons that fit neither category, referred to as miscellaneous. In the meta-analysis, effect sizes of the three larger sets of phonics types were compared.

In the database were seven phonics programs whose effectiveness was assessed in at least three different treatment-control group comparisons. All but one of the programs, Lovett's analogy program, taught synthetic phonics. These programs together with the dates of publication are listed below:

- Direct Instruction, also referred to as DISTAR and Reading Mastery (1969, 1978, 1979, 1980, 1987, 1988)

- Lovett's adaptation of Direct Instruction (1994)

- Lovett's adaptation of the Benchmark Word Identification program (1994)

- The Lippincott Basic Reading program (1963, 1981)

- Beck and Mitroff's New Primary Grades Reading System (1972)

- Orton Gillingham programs (1940, 1956, 1969, 1979, 1984)

- Sing, Spell, Read, and Write (1972).

For each program, there were at least three treatment-control group comparisons testing effects of that form of phonics instruction; so, effect sizes were examined separately in a meta-analysis. Most of these programs were developed over 20 years ago, providing researchers with more time to study them than recently developed programs. The question addressed in the meta-analysis was whether these programs were effective in promoting growth in reading and whether they differed in effectiveness. There was no apriori reason to expect any differences. Likewise there was no reason to expect these programs to be more effective than programs not in the set being compared.

Grade and Reading Ability

A question of particular interest to the Panel was when should phonics instruction begin. Should it be introduced in kindergarten when children may know very little about letters, phonemic awareness, or should it be started in 1st grade after children have received prereading or emergent reading experiences in kindergarten? According to Chall (1996a, b), beginners need to develop foundational knowledge such as concepts about print, phonological awareness, and letter names prior to formal reading instruction. Studies indicate that knowing letters and having phonemic awareness are essential for learning to use the alphabetic system to read and spell words (see the NRP review of phonemic awareness instruction). Thus, formal, systematic phonics instruction that expects students to learn to decode words in kindergarten may be too much.

On the other hand, in countries such as New Zealand and the United Kingdom, the practice of introducing children to reading and writing at the age of 5 in full-day programs has existed for many years. The Reading Recovery© program (Clay, 1993) is designed to pick up the stragglers having difficulty at the age of 6, when North American children are typically just beginning reading instruction. Thus, the notion that kindergartners are not ready for formal reading instruction at age 5 is questionable.

In some studies in the database, a middle road was taken. Children were introduced to simplified reading and spelling activities using a basic set of letters and sounds that they were taught. Instruction began by providing a foundation for students and then building on this to ease students into reading when they became ready for it. (See Blachman et al., 1999; Vandervelden & Siegel, 1997). In the meta-analysis, the contribution of phonics instruction at the kindergarten level was examined across studies that varied in how much phonics material was covered.

The most important grade for teaching phonics is thought to be 1st grade when formal instruction in reading typically begins in the United States. Children have foundational knowledge and are ready to put it to use in learning to read and write. In contrast, introducing phonics instruction in grades above 1st means that children who were taught to read in some other way may be required to switch gears in order to incorporate phonics procedures into their reading and writing. The database included studies that introduced phonics to students at various grade levels. The question addressed in the meta-analysis was whether the grade level in which phonics instruction was introduced made any difference in the outcomes observed. Another related question is whether phonics

instruction that was started in kindergarten is more effective than phonics instruction begun in 1st or 2nd grade. Data were probed for an answer to this question as well.

Phonics instruction has also been widely regarded as particularly beneficial to children with reading problems (e.g., Foorman et al., 1998). Many studies have shown that reading disabled children have exceptional difficulty decoding words (Rack, Snowling, & Olson, 1992). In fact, their level of performance falls below that of younger non-disabled readers who read at the same grade-equivalent level, indicating a serious deficit in decoding skill. Phonics instruction that teaches disabled readers to decode words should remediate this deficit and should enable these students to make better progress in learning to read. The meta-analysis evaluated the contribution made by phonics instruction to growth in reading among children having difficulty learning to read.

Two types of children with reading problems have been distinguished by researchers, children who are *unexpectedly* poor readers because their intelligence (an index of learning aptitude for some academic skills) is *higher* than their reading ability, and children whose below-average reading is not surprising given that their intelligence is also below average. Various labels such as dyslexic or learning disabled or reading disabled have been applied to children whose higher IQs are discrepant with their poor reading skill. Children whose lower reading scores are consistent with their lower IQs have been referred to as low achievers or garden variety poor readers (Stanovich, 1986). The question of interest was whether phonics instruction helps to remediate reading difficulties for both types of poor readers. Studies in the database were brought to bear on this question.

Delivery Systems for Teaching Phonics

There are various delivery systems that might be used to teach phonics. Tutoring one-on-one is regarded as the ideal form of instruction for students who are having difficulties because it allows teachers to tailor lessons to address individual students' needs. One of the best known tutoring programs is Reading Recovery© (Clay, 1993). The database included three studies that modified Reading Recovery© lessons to include systematic phonics instruction (Greaney et al., 1997; Santa & Hoien, 1999; Tunmer & Hoover, 1993). A total

of eight studies taught phonics through tutoring. The remainder of the studies utilized small groups or whole classes to deliver instruction. Of interest was whether one type of delivery system produced greater gains in reading than the other types. In the Panel's analysis of phonemic awareness training effects, comparison of instructional units revealed that small groups produced superior learning. However, it was expected that tutoring would be the most effective way to teach phonics.

Word Reading Processes: Assessing Growth

It is important to distinguish between the *methods* of teaching reading and the *processes* that learners acquire as they receive instruction and learn to read. Sometimes the two may be confused. For example, the term "sight word" has a "methods" meaning and a "process" meaning. As a method, sight words are the high-frequency, irregularly spelled words students are taught to read as unanalyzed wholes, often on flash cards, for example, said, once, their, come. In contrast, the "process meaning" of sight words refers to words that are stored in readers' heads and that enable them to read those words immediately upon seeing them. Not just high-frequency words but all words that readers practice reading become retained as sight words in memory.

Methods of teaching reading are aimed at helping learners acquire the processes they need to develop skill as readers. In considering how phonics instruction promotes growth in reading, it is important to describe the reading processes that learners are expected to acquire.

Learning to read can be analyzed as involving two basic processes (Gough & Tunmer, 1986; Hoover & Gough, 1990). One process involves learning to convert the letters into recognizable words. The other involves comprehending the meaning of the print. When children attain reading skill, they learn to perform both of these processes so that their attention and thought are focused on the meaning of the text while word reading processes operate unobtrusively and out of awareness for the most part. Children acquire comprehension skill in the course of learning to speak. Comprehension processes that children use to understand spoken language are thought to be the same ones that they use

to read and understand text. In contrast, children do not acquire word reading skill in the course of learning to speak. This achievement requires special experiences and instruction.

Many mental processes are active when readers read and understand text. Readers draw on their knowledge of language to create sentences out of word sequences. They access their background knowledge to construct meaning from the text. They retain this information in memory and update it as they interpret more text. Readers monitor their comprehension to verify that the information makes sense.

A central part of text processing involves reading the words. Four different ways can be distinguished (Ehri, 1991, 1994):

1. Decoding: Readers convert letters into sounds and blend them to form recognizable words; the letters might be individual letters, or digraphs such as TH, SH, OI, or phonograms such as ER, IGH, OW, or spellings of common rimes such as -AP, -OT, -ICK. Ability to convert letter subunits into sounds comes from readers' knowledge of the alphabetic system.

2. Sight: Readers retrieve words they have already learned to read from memory.

3. Analogy: Readers access in memory words they have already learned and use parts of the spellings to read new words having the same spellings (e.g., using -ottle in bottle to read throttle).

4. Prediction: Readers use context cues, their linguistic and background knowledge, and memory for the text to anticipate or guess the identities of unknown words.

Text reading is easiest when readers have learned to read most of the words in the text automatically by sight because little attention or effort is required to process the words. When written words are unfamiliar, readers may decode them or read them by analogy or predict the words, but these steps take added time and shift attention at least momentarily from the meaning of text to figuring out the words.

Readers need to learn how to read words in the various ways to develop reading skill. The primary way to build a sight vocabulary is to apply decoding or analogizing strategies to read unfamiliar words. These ways of reading words help the words to become familiar.

Processing letter-sound relations in the words through decoding or analogizing creates alphabetic connections that establish the words in memory as sight words (Ehri, 1992; Share, 1995).

Systematic phonics instruction is thought to contribute to the process of learning to read words in these various ways by teaching readers use of the alphabetic system. Alphabetic knowledge is needed to decode words, to retain sight words in memory, and to call on sight word memory to read words by analogy. In addition, the process of predicting words from context benefits from alphabetic knowledge. Word prediction is made more accurate when readers can combine context cues with letter-sound cues in guessing unfamiliar words in text (Tunmer & Chapman, 1998).

One purpose of the meta-analysis was to examine whether phonics instruction improves readers' ability to decode words and to read words by sight. To study the impact of phonics instruction on the various ways to read words, different measures have been used. The ability to decode words is tested by giving children regularly spelled words to read. The ability to decode novel words never read before is tested by having children read pseudowords. Children's sight vocabulary is examined by giving them miscellaneous words including irregularly spelled words that are ordered by grade level from preprimer to the highest grades.

Methodology

Database

An electronic search was conducted in two databases, ERIC and PsycINFO. Three sets of terms were used in the search. These terms were derived by the Panel on the basis of analyses of various reference guides including the Literacy Dictionary (Harris & Hodges, 1995), the Handbook of Research on Teaching the English Language Arts (Flood, Jensen, Lapp, & Squire, 1991), the Encyclopedia of English Studies and the Language Arts (Purves, 1994), and the Handbook of Reading Research (Barr, Kamil, Mosenthal, & Pearson, 1991; Pearson, Barr, Kamil, & Mosenthal, 1984).

-   **Set 1:** Alphabetic code, analogy approach, code emphasis, compare-contrast, decodable text, decoding, phonemic decoding, phonetic decoding, phonological decoding, direct code, direct instruction, Reading Mastery, explicit instruction,

explicit phonological processes, grapheme-phoneme correspondences, graphophonic, Initial Teaching Alphabet, letter training, letter-sound correspondences, linguistic method, McCracken, Orton-Gillingham, phoneme analysis, phoneme blending, phoneme-grapheme correspondences, phonics, Alphabetic phonics, analytic phonics, embedded phonics, structured phonics, synthetic phonics, systematic phonics, phonological processing, Recipe for Reading, recoding, phonological recoding, Slingerland approach, Spaulding approach, word study, word sort, words by analogy. These were combined using "or" statements, meaning that all articles indexed by any of these terms would be located.

- **Set 2:** Beginning reading, beginning reading instruction, instruction, intervention, learning to decode, reading improvement, reading instruction, remedial training, remedial reading, remediation, teaching, training, disabled readers, dyslexia, reading difficulties, reading disability, reading failure, reading problems. These were combined in the search using "or" statements.

- **Set 3:** Miscues, oral reading, reading ability, reading achievement, reading acquisition, reading aloud, reading comprehension, reading development, reading processes, reading skills, silent reading, story reading, word attack, word identification, word recognition, word reading, nonword reading. These, too, were combined with "or" statements.

The three sets of terms were used to locate potentially relevant studies in the two databases. Articles selected were those that included at least one term from each set. Because the term <u>spelling</u> had not been included in Set 1, the search was run a second time with spelling crossed with Set 2 and Set 3 terms. The first search uncovered 391 articles in PsycINFO and 520 articles in ERIC. The second search uncovered 252 articles in PsycINFO and 210 articles in ERIC. Abstracts were printed and screened.

To qualify for the analysis, studies had to meet the following criteria:

1. Studies had to adopt an experimental or quasi-experimental design with a control group.

2. Studies had to appear in a refereed journal after 1970.

3. Studies had to provide data testing the hypothesis that systematic phonics instruction improves reading performance more than instruction providing unsystematic phonics or no phonics instruction. To be considered an instance of phonics instruction, the treatment had to teach children to identify or use symbol-sound correspondences systematically.

4. Studies had to measure reading as an outcome.

5. Studies had to report statistics permitting the calculation or estimation of effect sizes.

6. Studies were not those already included in the National Reading Panel's meta-analysis of phonemic awareness training studies.

From the various lists of references, 75 studies that appeared to meet the criteria were identified and located. The goal was to analyze studies that resembled each other so that the corpus would be more homogeneous. Studies of instructional interventions that might be found in schools were sought. Short-term laboratory studies and studies that provided instruction on only a limited set of processes were eliminated. Also eliminated were studies that simply compared different forms of phonics instruction but did not include a control group receiving reduced phonics or no phonics. Of the 75 studies screened, 38 were retained and 37 were eliminated from the final set used to calculate effect sizes. The reasons for eliminating studies and the numbers of studies eliminated are listed in Table 1 on the next page.

Some minor deviations from the above procedures occurred. More recent studies that would not yet have appeared in electronic searches were obtained from current issues of journals and preprints of in press papers sent to members of the Panel. Also, Blachman et al. (1999) conducted a 3-year longitudinal study to evaluate the effects of phonemic awareness and phonics instruction on children as they progressed from kindergarten through 2nd grade. Results of the first year were published as a separate study and included in the Panel's phonemic awareness meta-analysis. Results of the more extensive 3-year study were included in the phonics instruction database. This was the only study analyzed in both reports.

Table 1

Reasons for Excluding Studies From the Database

BASIS FOR REJECTION	NUMBER
Control group missing or inadequate:	5 studies
Short-term, focused too limited, or laboratory study:	14 studies
Inadequate statistics:	8 studies
Inadequate outcome measures:	3 studies
Not a study of phonics instruction:	2 studies
Duplicate data reported in another publication already considered:	5 studies
Total:	37 studies

The primary statistic used in the analysis of performance on outcome measures was effect size, indicating whether and by how much performance of the treatment group exceeded performance of the control group, with the difference expressed in standard deviation units. The formula used to calculate raw effect sizes for each treatment-control comparison consisted of the mean of the treatment group minus the mean of the control group divided by a pooled standard deviation.

From the 38 studies entered into the database, 66 treatment-control group comparisons were derived. There were six cases in which the same control group was compared to two different phonics treatment groups. There was one study in which the same control group was compared to four different treatments (Lovett et al., in press). Each comparison was treated as a separate case with separate effect sizes in the database.

Studies were coded for several characteristics that were included as moderators in the meta-analysis:

- Type of phonics program (synthetic vs. larger subunits vs. a combination of synthetic and larger subunits vs. miscellaneous)

- Specific phonics program if replicated in at least three comparisons

- Type of control group (basal, regular instruction, whole language, whole word, miscellaneous)

- Group assignment procedure (random assignment or nonequivalent groups)

- Number of participants (blocked into quartiles)

- Grade level or age

- Reading ability (normally developing, at risk/low achiever, reading disabled)

- Socioeconomic status (low, middle, varied, not given)

- Instructional delivery unit (class, small groups, 1:1 tutoring).

The studies, their properties, and effect sizes are listed in Appendix G.

Although the length of treatment was coded, it was not used as a moderator variable. Many of the studies were vague about the amount of time devoted to phonics instruction; so, it was not possible to calculate precise amounts of time spent, particularly in classroom studies which provided instruction regularly throughout the school year. Also, treatment length was confounded with other variables considered to be more important,

such as whether students were tutored or taught in classes, whether students were poor or normally developing readers, whether students were beginners or older readers when they began instruction.

Some studies in the database selected normally developing readers to include in their experiments whereas other studies singled out poor readers. These students were grouped into four types of readers for analysis:

1. Normally developing readers: this category included studies in which poor readers were excluded and studies where no attempt was made to distinguish children by reading ability.

2. Disabled readers: this category included children who were identified as reading disabled according to IQ-reading discrepancy criteria in standard use by researchers, or were given tests to determine that the disability was reading-specific; in some cases, exclusionary criteria were applied as well (e.g., no neurological, behavioral, economic, or emotional disorders); most of these children were above 1st grade.

3. Children at risk for developing reading difficulties in the future (kindergartners and 1st graders).

4. Children who were below average in their reading referred to as low achievers (children above 1st grade).

The latter two groups included children who exhibited poor letter knowledge, poor phonemic awareness, or poor reading skills, or those in schools with low achievement, or those identified by teachers as needing special help in reading, or those who qualified for remedial programs in schools but the criteria for selection were not specified. The at-risk label was applied to children in kindergarten and 1st grade because they were still at a beginning level in their learning. Children labeled low achievers in reading were those in 2nd grade and above whose identity as poor readers was considered to be better established. Both groups included children who also had lower than average IQs qualifying them as garden variety poor readers with generally low academic achievement, but the groups were not limited to children with low IQs because researchers either did not measure IQ or did not use it to limit the readers selected for study.

Six types of outcomes assessing growth in reading or spelling were distinguished:

- Decoding of real words chosen to contain regular spelling-to-sound relationships

- Reading nonsense words or pseudowords chosen to represent regular spelling-to-sound relationships.

- Word identification (in some cases, words were chosen to represent irregular spelling-to-sound relationships)

- Spelling, assessed using either developmental stages for younger children (Bear et al., 2000) or number of words correct

- Comprehension of material read silently or orally

- Oral reading of connected text (accuracy).

Measures reported in studies were classified into these types, and effect sizes were computed for each type of outcome. Some studies included several measures of an outcome type and reported means on each measure. In these cases, effect sizes were calculated on each measure and then averaged. This step insured that no single treatment-control comparison contributed more than one effect size to any single outcome category. Some studies included tests to assess whether students were able to read or spell words that were taught directly during phonics instruction. These results were not included as outcomes in the database.

For each comparison, the mean effect size was calculated across whichever of the six measures had been assessed in that study. This yielded an overall outcome measure for each comparison. When studies reported performance on a general reading test but no more specific tests, the overall effect size was based on the general measure. Outcomes that did not fit into the above categories were not entered into the database.

Performance of students was measured at various points before, during, and after instruction. Entered into the database were outcomes of posttests measured at three points in time: at the end of training, at the end of the first school year if the program was taught for more than one year, after a delay following training to assess long-term effects. The type of posttest most commonly given was that occurring at the end of the program or at the end of the school year when the program continued; so, this was the outcome used in most of the analyses of moderator variables.

In the categorization of outcome measures, no distinction was drawn between standardized and experimenter-devised tests. Comprehension measures tended to be standardized. Oral reading measures tended to be informal reading inventories that were neither standardized nor developed specifically for the study. Word lists were both standardized and experimenter-devised. Standardized tests of word reading most commonly came from the Woodcock Johnson Achievement series, the Woodcock Reading Mastery Test, and the Wide Range Achievement (WRAT) test. In general, standardized measures tend to produce smaller effect sizes than experimenter-devised measures. This was observed in the NRP's analysis of effects of phonemic awareness instruction on measures of word reading and spelling. One reason is that standardized tests are designed to assess reading across a wide range of ability levels and hence are less sensitive to differences at any one level in the range. Thus, aggregating the two types of tests would be expected to underestimate effect sizes slightly.

The information and statistics required to generate and analyze effect sizes were entered into a separate database using Microsoft Excel and SPSS. The data entered included identification of the study, codes for the information listed above, means and standard deviations of treatment and control groups on outcome measures, pooled standard deviations, raw effect sizes (g) and effect sizes weighted for the size of the sample (d). When means and standard deviations were not available in the article, DSTAT was used to estimate effect sizes based on t or F values. When pretest differences between treatment and control groups were reported, effect sizes were calculated to eliminate these differences as far as possible.

The DSTAT statistical package (Johnson, 1989) was employed to calculate effect sizes and to test the influence of moderator variables on effect sizes. Each moderator variable had at least two levels. Tests were conducted to determine whether the mean weighted effect size (d) at each level was significantly greater than zero at $p < 0.05$, whether the individual effect sizes at each level were homogeneous ($p < 0.05$), and whether effect sizes differed significantly at different levels of the moderator variables ($p < 0.05$).

Consistency With the Methodology of the National Reading Panel

The methodology approved by the National Reading Panel was adopted. The search was conducted in accordance with most of the prescribed procedures. Studies that were not published in peer-reviewed journals were excluded. All of the studies in the data base utilized experimental or quasi-experimental designs. (Studies using a multiple baseline design were not included.) The studies were coded for most of the specified categories plus some additional categories of interest for this particular analysis. Properties left uncoded were those where information was rarely provided. More properties were coded than were considered in the analysis. One reason for not analyzing effects of moderator (coded) variables on outcomes was that there were insufficient numbers of comparisons to provide a valid analysis of these effects.

The Panel determined that a meaningful meta-analysis could be conducted on the data. The means and standard deviations that were used to calculate effect sizes were verified by checking all of them at least twice. Intercoder reliability was conducted on the variables used in the meta-analysis and exceeded the prescribed level of 90%. Disagreements were resolved by discussion and consensus.

Results

Characteristics of Studies in the Data Set

There were 38 studies from which 66 treatment-control group comparisons were derived. Each comparison could contribute a maximum of six effect sizes, one per outcome measure. However, few studies included measures of all the outcomes. The most commonly assessed outcome (i.e., at the end of training or at the end of one year, whichever came first) was word identification consisting of 59 effect sizes. The least common outcome was oral reading with 16 effect sizes. The other outcomes ranged from 30 to 40 effect sizes. Whereas 76% of the effect sizes involved reading or spelling single words, only 24% involved text reading. Although there is a marked imbalance favoring single words, this is not surprising given that phonics instruction is aimed primarily at improving children's ability to read and spell words.

Many of the studies limited instructional attention to children with reading problems. These studies accounted for 65% of the comparisons, with 38% involving poor readers considered "at risk" or low achieving, and 27% involving children diagnosed as reading disabled (RD). Studies involving 1st graders were overrepresented in the database compared to other grades and accounted for 38% of the comparisons. Fewer studies involved kindergartners and children in 2nd through 6th grades, with these groups contributing 12% and 23% of the comparisons, respectively. Children in the RD group spanned several ages and grades, ranging from ages 6 to 13 and grades 2 to 6. Several properties of the studies in our database were examined. Of interest was whether the studies were older or more recent. A tally revealed the following distribution:

> 1970 to 1979: 1 study
>
> 1980 to 1989: 9 studies
>
> 1990 to 2000: 28 studies

Thus, the majority of the studies were conducted over the last 10 years. Most (66%) were carried out in the United States, but 24% were done in Canada, and the remainder in the United Kingdom, Australia, and New Zealand. Thus, the evidence came from a variety of locales. Other properties of comparisons in the database are listed in Table 2 in Appendix D.

Effects of Phonics Instruction on Outcome Measures

The statistic used to assess the effectiveness of phonics instruction on children's growth in reading was effect size which measures how much the mean of the phonics group exceeded the mean of the control group in standard deviation units. An effect size of 1.0 indicates that the treatment group mean was one standard deviation higher than the control group mean, suggesting a strong effect of training. An effect size of 0 indicates that treatment and control group means were identical, suggesting that training had no effect. To judge the strength of an effect size, values suggested by Cohen (1988) are commonly used. An effect size of 0.20 is considered small; a moderate effect size is 0.50; an effect size of 0.80 or above is large.

An overall effect size was calculated for each of the 66 treatment-control group comparisons. This was the average of the six specific outcome effect sizes (i.e., decoding, word reading, comprehension, etc.) or the effect size from a general reading measure if no specific outcomes were measured. In the analyses, this overall effect size is interpreted as assessing the impact of phonics instruction on growth in reading. Although one of the six was a spelling measure, spelling effect sizes contributed only 16% of the effect sizes that were averaged and reading measures contributed the rest (84%). Mean effect sizes obtained on various outcomes associated with levels of the moderator variables are reported in Table 3 (Appendix E). Effect sizes were tested statistically to determine whether each was significantly greater than zero, indicating that superior performance of phonics-trained groups over control groups was not a result of chance at $p < 0.05$.

Inspection across the effect sizes listed in Table 3 reveals that the vast majority were significantly greater than zero (those marked with an asterisk). This indicates that systematic phonics instruction was effective across a variety of conditions and characteristics. The overall mean effect size of phonics instruction on reading was $d = 0.41$ when effects of programs were tested at their conclusion. A few programs lasted longer than 1 school year. To obtain another index of effects, outcomes measured either at the end of the program or the end of the first school year, whichever came first, were calculated. Results revealed an effect size of $d = 0.44$. These findings indicate that the effect produced by phonics instruction on reading was moderate in size. Unless otherwise stated, the test point used to assess effects of moderator variables in the meta-analyses was that occurring at the end of training or at the end of the first school year, whichever came first.

Phonics instruction in most of the studies lasted 1 school year or less. However, there were four treatment-control comparisons in which longer training was provided. In these studies, children at risk for reading problems began phonics instruction in kindergarten or 1st grade and continued for 2 or 3 years. Outcomes were measured at the end of each school year (Blachman et al., 1999; Brown & Felton, 1990; Torgesen et al., 1999). Characteristics and results of the four comparisons drawn from these studies are presented in Table 4. Mean effect sizes across the four

comparisons were sizeable and their strength was maintained across the grades: kindergarten $\underline{d}$ = 0.46; 1st grade $\underline{d}$ = 0.54; 2nd grade $\underline{d}$ = 0.43. This indicates the value of starting phonics early and continuing to teach it for 2 to 3 years. (See results below for additional evidence regarding the value of teaching phonics early.) In the Blachman et al. (1999) study, instruction was not given to all 2nd graders but only to those who had not attained the goals of the program after 2 years of instruction. These findings point to the importance of programs providing tests for teachers to use to determine which children need additional systematic phonics instruction and which have mastered the processes taught.

A few studies examined effects of phonics instruction several months after the treatment had ended. The specific comparisons together with their properties are listed in Table 4 (Appendix E). Followup tests were administered from 4 months to 1 year after training. As shown in Table 3, the effect size remained significantly greater than zero, indicating that the impact of phonics instruction lasted well beyond the end of training although its size was somewhat diminished (from $\underline{d}$ = 0.51 to $\underline{d}$ = 0.27).

The aim of phonics instruction is to help children acquire knowledge and use of the alphabetic system to read and spell words. Phonics was expected to exert its greatest impact on the ability to decode regularly spelled words and nonwords. Phonics instruction was also expected to exert a large effect when spelling was measured using a developmental spelling scale, which gives credit for letter-sound spellings as well as correct spellings (e.g., Bear et al., 2000; Blachman et al., 1999). These capabilities all benefit directly from alphabetic knowledge. Phonics instruction was expected to exert a significant but smaller impact on the ability to read miscellaneous words that included irregularly spelled words. Although alphabetic knowledge is not helpful for decoding irregularly spelled words, it does help children remember how to read these words (Ehri, 1998). Phonics instruction was expected to impact text reading processes. The effect was expected to be significant but smaller because its influence is indirect.

From Table 3 (Appendix E), it is apparent that effect sizes for all six types of measures were statistically greater than zero, indicating that phonics instruction significantly improved performance on all of the outcome measures examined, not only word reading and spelling but also text processing. Inspection of the size of the effects provided support for the various hypotheses. The strongest effects occurred on measures of decoding regularly spelled words ($\underline{d}$ = 0.67) and pseudowords ($\underline{d}$ = 0.60). These effects were statistically larger than effects observed on the other measures which did not differ from each other. This indicates that phonics instruction was especially effective in teaching children to decode novel words, one of the main goals of phonics.

Effect sizes on comprehension measures ($\underline{d}$ = 0.27) and oral reading measures ($\underline{d}$ = 0.25) were statistically greater than zero, indicating that phonics instruction significantly improved children's text processing skills as well as their word reading skills. The fact that effects of phonics instruction on reading comprehension were positive serves to dispel any belief that teaching phonics to children interferes with their ability to read and comprehend text. Quite the opposite is the case.

Several reasons explain why effects were somewhat smaller on text processing measures than on word reading measures. The tests of comprehension were predominantly standardized tests which are less sensitive when the range of performance is limited. The target of phonics instruction is teaching children how to read words. Although word recognition skill influences how well children can read and comprehend text, there are other processes that are important as well. Moreover, readers can still get meaning from text even when they cannot read some of the words.

Analysis of Moderator Variables

Studies in the database varied in several respects that were coded and analyzed as moderator variables. Of interest was whether these moderator variables enhanced or limited the effectiveness of systematic phonics instruction on growth in reading. It is important to recognize the limitations of this type of analysis and the tentative nature of any conclusions that are drawn. Findings involving the impact of moderator variables on effect sizes cannot support strong claims about moderators being the cause of the difference. Moderator findings are no more than correlational. The biggest source of uncertainty is whether there is a hidden variable that is confounded with the moderator and is the true cause of the difference.

Characteristics of Students

The students who received phonics instruction across the studies varied in two important ways that were expected to make a difference on the effect sizes produced by phonics instruction: their age or grade in school, and their reading ability. Kindergartners, particularly those at risk, know little about letters and sounds. Typically they are nonreaders. For them, phonics instruction begins by teaching letter shapes, letter sounds, phonemic awareness, and how to apply these in simplified reading and writing tasks. Later in kindergarten or at the beginning of 1st grade, formal reading instruction begins with much ground to cover. Children typically start as emergent readers and by the end of 1st grade are able to read text independently. In systematic phonics programs, extensive instruction is provided to develop children's knowledge of the alphabetic system and how to use this knowledge to read words in and out of text. The greatest impact of phonics instruction is expected to occur in helping 1st graders get off the ground in learning to read.

Designers of phonics programs to teach beginning reading expect children to start receiving instruction in their programs when the children are in kindergarten or 1st grade before they have acquired any reading skill. Programs are designed so that children usually continue receiving instruction at least through 2nd grade. What happens when these programs are taught to children above 1st grade who have already acquired some reading skill with some other program is less clear. Are the older children given 1st grade catch-up instruction? Do the phonics strategies that they are taught compete or conflict with the reading skills and strategies that they have already acquired? If so, what is done about this instructionally? There are many uncertainties surrounding the introduction of phonics instruction to children in the upper grades who have already moved into reading.

The database that the Panel analyzed included several studies with older children beyond 1st grade. Many of these studies involved disabled readers or low achieving readers who received remedial instruction designed to address the problems of poor readers. However, there were also a few studies in which phonics instruction was provided to normally developing readers who had already received instruction in other unspecified programs in the earlier grades. It is important to recognize that the question addressed in the meta-analysis of these studies was whether introducing phonics instruction presumably as a new program for these older children was effective in promoting their growth in reading.

Younger vs. Older Children

To analyze the impact of age and grade combined, two groups of children were distinguished: the younger children in kindergarten and 1st grade; and the older students in 2nd through 6th grades. The latter group included the mixed age/grade comparisons involving reading disabled (RD) children and low achieving readers. The outcome variable was the effect sizes on the immediate posttest given either at the end of training or at the end of the first year of the program, whichever came first.

From Table 3 (Appendix E), it is apparent that systematic phonics instruction produced a significant impact on children's growth as readers in both groups, as indicated by effect sizes statistically greater than zero. However, phonics instruction made a larger contribution to younger children's growth as readers ($\underline{d}$ = 0.55) than to older children's growth ($\underline{d}$ = 0.27). The difference in effect sizes favoring younger children was statistically significant.

The pool of effect sizes among the younger students was not homogeneous; so, effects were examined separately for kindergartners and 1st graders. From Table 2, it is evident that effect sizes were very similar, $\underline{d}$ = 0.56 for kindergartners and $\underline{d}$ = 0.54 for 1st graders. This shows that a moderate and significant effect size typified children in both grades. According to Chall (1992), phonics instruction should exert its greatest impact in the early grades. These findings show that effects were equally strong in both kindergarten and 1st grade, indicating that "early" includes both of these grades. There were many more studies of the impact of phonics in 1st grade than in kindergarten, so the 1st grade findings are more reliable than the kindergarten findings.

Whereas the database on phonics instruction included only seven comparisons involving kindergartners, the National Reading Panel's database of phonemic awareness training studies included 40 kindergarten

comparisons that measured reading as an outcome. In the PA analysis, effects were moderate in size and statistically significant. The effect size in the PA analysis ($\underline{d} = 0.48$) was close to the effect size produced by phonics instruction ($\underline{d} = 0.58$). Combined, these findings clearly support the importance of teaching phonemic awareness and grade-appropriate phonics in kindergarten. Indeed, some of the phonemic awareness training studies that taught children to analyze phonemes using letters would have qualified as phonics studies. If these PA studies had not been excluded from the phonics database, there would have been more kindergarten comparisons.

The above findings suggest that when phonics instruction is introduced and taught in kindergarten or 1st grade to readers who have little reading ability, it produces a larger effect than when phonics is introduced in grades above 1st grade with readers who have already acquired some reading skills. However, before concluding that phonics is truly less effective with older children, it is important to consider several mitigating factors. The majority of the comparisons in the older group, 78%, involved either low achieving or disabled readers. Remediating their reading problems may be especially difficult. In addition, there were only seven comparisons involving older, normally developing readers, and four of these came from one study using the Orton-Gillingham method, a program developed for disabled readers, not for non-disabled upper elementary level readers. Perhaps other types of phonics programs designed expressly to improve reading in older non-disabled children might prove more effective. This question awaits more research.

The set of effect sizes for the older students proved to be homogeneous, indicating that chance, rather than other moderator variables, explains the variation in effect sizes. The two types of poor readers, low achievers and RDs, contributed the majority of the effect sizes to this pool. These findings indicate that low achieving readers and disabled readers do not differ in their response to phonics instruction.

Specific Outcomes in Younger Readers

Because the younger and older children differed in their response to phonics instruction, the question of whether phonics instruction impacted children's ability to decode and spell words and to read text was answered separately for the two groups. Results in Table 3

(Appendix E) show that, among kindergartners and 1st graders, phonics instruction produced significant growth on all six outcome measures whose effect sizes were statistically greater than zero. Because a central goal in phonics programs is to teach students to decode novel words, one would expect the strongest effects to be evident in decoding tasks. This is what was found. The largest effect size was produced on the measure of decoding regularly spelled words ($\underline{d} = 0.98$). Moderately large effects were also produced on measures of decoding pseudowords ($\underline{d} = 0.67$) and spelling words ($\underline{d} = 0.67$). The effect size was somewhat reduced on the word identification outcome ($\underline{d} = 0.45$). This is not surprising since tests of word identification often included irregularly spelled words not amenable to decoding.

Phonics instruction with its emphasis on teaching letter-sound relations would be expected to improve beginning readers' ability to spell words by writing the sounds they hear. Studies with younger children commonly employed developmental spelling scoring systems that gave credit for phonetically plausible spellings, for example, spelling <u>feet</u> as FET or <u>car</u> as KR (Tangel & Blachman, 1995; Morris & Perney, 1984). This may explain the sizeable effect observed on the spelling outcome ($\underline{d} = 0.67$).

Among beginning readers, phonics instruction exerted a significant impact on reading comprehension. The effect size, based on ten 1st grade and one kindergarten comparisons, was moderate ($\underline{d} = 0.51$). However, the effect size on another measure of text reading, oral reading, was smaller but also significantly greater than zero ($\underline{d} = 0.23$ based on two kindergarten and four 1st grade comparisons). Why phonics skills facilitated reading comprehension more than oral reading is not clear. It may have to do with the nature of the tests. Standardized comprehension tests at this level generally use extremely short (usually one sentence) "passages." On these short passages, the effects of decoding should be strong. Some tests, such as the Gates-MacGinitie, favor phonetically regular words in these passages. Oral reading measures, on the other hand, use longer passages, sometimes containing pictures which would enhance the utility of context.

One would expect effect sizes on text reading and word reading to be similar because 1st graders' ability to read and understand text is heavily influenced by their ability to read the words in the text, perhaps somewhat more so than in later grades. This is supported by Juel (1994) who found a very high correlation between word recognition and reading comprehension in 1st grade ($\underline{r}$ = 0.87) and found that the correlation was somewhat lower in 2nd grade ($\underline{r}$ = 0.73).

In sum, these findings show that systematic phonics instruction helped beginning readers acquire and use the alphabetic system to read and spell words in and out of text. Children who were taught phonics systematically benefited significantly more than beginners who did not receive phonics instruction in their ability to decode regularly spelled words and nonwords, in their ability to remember how to read irregularly spelled words, and in their ability to invent phonetically plausible spellings of words. In addition, phonics instruction contributed substantially to students' growth in reading comprehension and somewhat less to their oral text reading skill.

Specific Outcomes in Older Readers

Students above the 1st grade were introduced to phonics instruction in their classes or in pull-out programs for periods lasting up to a school year. These students included children who were low achieving readers as well as children diagnosed as reading disabled. Effects of phonics instruction on six outcome measures were compared. Results in Table 3 (Appendix E) show that substantial growth occurred in learning to decode regularly spelled words ($\underline{d}$ = 0.49) and pseudowords ($\underline{d}$ = 0.52), with effect sizes statistically greater than zero in the moderate range. This shows that phonics programs were significantly more effective than control programs in improving these students' knowledge and use of the alphabetic system which is the focus of phonics programs. Growth in the reading of miscellaneous words with irregularities was somewhat smaller but significant ($\underline{d}$ = 0.33), indicating that phonics improved students' ability to read irregularly spelled words, presumably by improving their memory for these words.

In contrast to strong positive effects of phonics instruction on measures of word reading, these programs were not more effective than other forms of instruction in producing growth in spelling ($\underline{d}$ = 0.09).

This effect size was not statistically different from zero. Likewise, phonics programs did not produce significant growth in reading comprehension ($\underline{d}$ = 0.12) although a small, statistically significant effect was observed on oral reading ($\underline{d}$ = 0.24).

Because the comparisons involving older children included a large number focusing on disabled readers, the 17 RD comparisons were analyzed separately. Effect sizes proved almost identical to those for the larger group reported in Table 3 (Appendix E) with one important exception. The effect size on the measure of reading comprehension, though small, was statistically greater than zero ($\underline{d}$ = 0.27, based on eight comparisons that were homogeneous). This indicates that, contrary to the general finding of no effect, systematic phonics instruction did help reading disabled students comprehend text more successfully than nonsystematic/no-phonics programs.

Because most of the comparisons above 1st grade involved poor readers (78%), the conclusions drawn about the effects of phonics instruction on specific reading outcomes pertain mainly to them. Findings indicate that phonics instruction helps poor readers in 2nd through 6th grades improve their word reading skills. However, phonics instruction appears to contribute only weakly, if at all, in helping poor readers apply these skills to read text and to spell words. There were insufficient data to draw any conclusions about the effects of phonics instruction with normally developing readers above 1st grade.

The absence of effects on spelling is noteworthy since the same finding was detected in the Panel's meta-analysis of phonemic awareness instruction. In the PA review, the Panel found that younger readers experienced growth in spelling as a result of phonemic awareness training, but the older disabled readers did not show improvement over controls. One possible explanation is that poor readers experience special difficulty learning to spell (Bruck, 1993). Remediation of this difficulty may require special instruction targeted at spelling. Another explanation may be that as readers move up in the grades, remembering the spellings of words is less a matter of applying letter-sound correspondences and more a matter of knowing more advanced spelling patterns and morphologically based regularities which is not typically addressed in phonics instruction.

Further research is needed to explore the value of phonics instruction in grades beyond 1st grade. Perhaps phonics instruction could be made stronger by combining it with instruction that helps children learn to read words in other ways, specifically, reading words from memory, reading words by analogy to known words, and reading words using spelling patterns and multisyllabic decoding strategies. Some phonics programs in the database did teach children about spelling patterns and the use of an analogy strategy to read words (see results presented below). Also it may be important for phonics programs to include systematic instruction in reading fluency and automaticity when phonics is taught to older students. A few of the programs in the database included exercises to promote fluency. Very likely, phonics programs that emphasize decoding exclusively and ignore the other processes involved in learning to read will not succeed in making every child a skilled reader.

Separation of Reader Ability Groups at Each Grade Level

To clarify whether and how readers with different reading abilities across the different grades responded to phonics instruction, treatment-control group comparisons were grouped by grade and reading ability. There were 62 comparisons with posttests administered when the program was completed or at the end of the first year of the program, whichever came first. Table 5 (Appendix E) shows how these comparisons were distributed across the grade-by-reader-ability cells.

Six groups were formed for the meta-analysis:

- 1st grade normally achieving readers

- 2nd through 6th grade normally achieving readers

- kindergarten children at risk for reading problems

- 1st grade children at risk

- 2nd through 6th grade low achievers

- disabled readers.

More precise grade and age information is given in Table 2 (Appendix D), which lists characteristics of each treatment-control group comparison.

The outcome measure was the overall effect size averaged across the six specific measures. Effect sizes significantly greater than zero were evident for five of the six groups of readers. From Table 3, it is apparent

that phonics instruction contributed to growth in reading in all groups but the 2nd through 6th grade low achiever group. Among the at-risk and normal readers in kindergarten and 1st grades, effect sizes were moderate to high, ranging from $\underline{d} = 0.48$ to $\underline{d} = 0.74$. Effect sizes were smaller for 2nd though 6th grade normal readers ($\underline{d} = 0.27$) and disabled readers ($\underline{d} = 0.32$). These findings extend the analysis above by revealing effect sizes for specific reader ability groups at each grade level. Findings indicate that the strong impact of phonics instruction was evident in normally developing 1st graders as well as at-risk kindergartners and 1st graders.

There was one group for whom phonics instruction failed to exert a statistically significant impact on the students' growth in reading. This occurred in the eight comparisons involving low achievers in 2nd through 6th grades ($\underline{d} = 0.15$). Although smaller, the effect size for low achievers did not differ significantly from the effect size of disabled readers ($\underline{d} = 0.32$).

Alternative explanations for the ineffectiveness of phonics instruction with older poor readers in 2nd through 6th grades can be offered. Their reading difficulties may have arisen from sources other than decoding, such as lack of fluency or poor reading comprehension skills (see other sections of the NRP report for elaboration of these reading processes). The fact that the IQs of some of the children in these studies were below normal points to comprehension difficulties as a possibility. Another explanation may be that these children were not given sufficiently intensive phonics instruction to remediate their difficulties. In Table 4 are listed properties of the treatment-control group comparisons involving low achievers. Inspection of the characteristics of these studies reveals that only one provided tutoring, thought to be the most effective way to teach phonics (but see below), whereas seven involved class instruction. However, there may be too few studies of low achieving readers in the database (only eight) to draw firm conclusions. Further research is needed to explore how best to remediate their reading difficulties.

Effects of Phonics Instruction Lasting 2 to 3 Years

The evidence on older readers above 1st grade reviewed so far provides no information about the effects of phonics instruction on older students who began phonics instruction in kindergarten or 1st grade.

However, there is relevant evidence in the database. For four comparisons, phonics instruction was introduced in kindergarten or 1st grade to at-risk readers and continued beyond 1 year (Blachman et al., (1999); Brown & Felton, 1990; Torgesen et al., 1999). These treatment-control group comparisons are listed in Table 4 (Appendix E). At the end of 2nd grade, after 2 to 3 years of instruction, the mean effect size was $\underline{d} = 0.43$. This is substantially higher than the mean effect size observed for older children receiving only 1 year of phonics instruction in grades beyond 1st ($\underline{d} = 0.27$). Because there are so few cases contributing effect sizes, the results are mainly suggestive. They suggest that when phonics instruction is taught to children at the outset of learning to read and continued for 2 to 3 years, the children experience significantly greater growth in reading at the end of training than children who receive phonics instruction for only 1 year after 1st grade.

SES

One additional characteristic of children was examined as a moderator variable, their socioeconomic status. Two different levels were represented in the database, low SES and middle SES. Also present were studies where SES was stated to vary and studies where it was not given. Table 3 shows that effect sizes were greater than zero in all cases. Phonics instruction exerted its strongest impact on low SES children ($\underline{d} = 0.66$). Its impact was somewhat less in middle SES students ($\underline{d} = 0.44$) although these two values did not differ statistically. These findings indicate that phonics instruction contributes to growth in reading in both low- and middle-class students.

Characteristics of Phonics Instruction

The treatment-control group comparisons were categorized by the type of systematic phonics instruction taught. In all studies, the programs were identified in sufficient detail to determine that systematic phonics was taught. However, some reports provided less description than others. For programs that were well known or were fully described, the Panel was able to make judgments about their characteristics and fit them into categories. Programs that were not described sufficiently were included in the miscellaneous category. (Publications describing programs are referenced in Appendix C.)

Types of Programs

It is important to recognize that the systematic phonics programs in the database varied not just in the way that the Panel categorized them but also in many other potentially important ways. However, the Panel's choice of categories was limited by the information provided in studies. Most authors mentioned whether the program emphasized synthetic phonics or the teaching of blending using larger subunits of words. However, other properties of programs were not consistently mentioned. Some especially important properties, such as the set of letter-sound relations covered were rarely mentioned. The four categories that were employed are listed in Table 2 (Appendix D) along with the specific treatment-control group comparisons in each category. (For the future, the Panel urges researchers to provide full descriptions of programs that are studied. Journal editors also should insist on this.)

Programs that emphasized systematic synthetic phonics were placed in one category. These programs taught students to transform letters into sounds (phonemes) and to blend the sounds to form recognizable words. This was by far the most common type of program, utilized in 39 of the comparisons. Some of the programs were developed by researchers while others were published programs, some widely used in schools, for example, Jolly Phonics, the Lindamood ADD program, the Lippincott program, Open Court, Orton Gillingham, Reading Mastery (also known as Direct Instruction or DISTAR), and Sing Spell Read & Write.

The second category of programs did not emphasize a synthetic approach at the phonemic level. Rather children were taught to analyze and blend larger subunits of words such as onsets, rimes, phonograms, or spelling patterns along with phonemes. Some of these programs were referred to as embedded code programs because grapheme-phoneme relations were taught in the context of words and text. Teaching children to segment and blend words using onsets and rimes taught them about units as small as graphemes and phonemes because onsets (i.e., the initial consonants in words) are very often single phonemes. In some programs, recognizing rimes in words provided the basis for teaching students the strategy of reading new words by analogy to known words sharing the same rimes. Words in texts were built from linguistic patterns. Writing

complemented reading in most programs. The programs in this category included Edmark, Hiebert's embedded code program, three Reading Recovery© programs modified to include systematic phonics, and a program derived from the Benchmark Word Identification program.

One of the 11 studies in the Larger Unit category, that by Tunmer and Hoover (1993), produced an atypical effect size, d = 3.71, which was much larger than the other effects. It should be noted that this study was atypical in that it was more intensive than most others. It involved one-on-one tutoring by highly trained teachers, and it combined phonemic awareness, phonics, and Reading Recovery© instructional strategies. To reduce the influence of this comparison on the overall mean, its effect size was reduced to equal the next largest effect size in the set, d = 1.41. (This method of adjusting effect sizes to deal with outliers was only applied in analyses that involved a small number of comparisons.)

The third category, referred to as miscellaneous, consisted of phonics programs that did not fit into the synthetic or larger unit categories. In some studies, the descriptions of programs did not state that a synthetic strategy was taught. If the program was not known to teach this decoding strategy, then it was placed in the miscellaneous category. Also, if the scope of instruction was limited and did not constitute a full phonics program (i.e., Haskell et al., 1992; Lovett et al., 1990), it was considered to be miscellaneous. This set included a spelling program, traditional phonics basal programs, and some researcher-devised instruction that focused on word analysis procedures.

The fourth category, referred to as combination programs, included only two comparisons. However, these could not be fit into the other categories because they examined the effects of teaching two of the other categories, a synthetic phonics program and a larger-units word analogy program (Lovett et al., in press). The comparisons differed in the order that the two programs were taught. The mean effect size for the combined programs was d = 0.42.

Effect sizes reported in Table 3 show that programs in all three categories produced effect sizes that were significantly greater than zero. This verifies that the three types of phonics programs were more effective than control programs in helping children learn to read. The 39 synthetic phonics programs produced a moderate impact on growth in reading (d = 0.45). The 11 programs that emphasized larger units created a somewhat smaller impact (d = 0.34) and likewise the ten miscellaneous programs' effect was smaller (d = 0.27). However, the three effect sizes did not differ statistically from each other (p > 0.05). There were relatively few comparisons in the larger unit group. Additional research would be useful for determining whether the small difference between the synthetic and larger unit approaches is a reliable one.

Specific Phonics Programs

There were seven phonics programs that were studied in three or more treatment-control comparisons. The identities of programs and properties of the comparisons testing their effectiveness are listed in Table 6 (Appendix F). Descriptions of the programs are provided in Table 7 (Appendix E). Effect sizes of these comparisons were subjected to a meta-analysis. Results in Table 3 (Appendix F) reveal that all effect sizes were statistically greater than zero, indicating that all the phonics programs produced significantly greater growth in reading than control group programs. The sets of effect sizes for all but one of the programs proved to be homogeneous. Effect sizes ranged from a high of d = 0.68 for the Lippincott program to a low of d = 0.23 for the Orton-Gillingham-based programs. Possible reasons for lower effect sizes in the case of Orton Gillingham comparisons are evident in Table 6 (Appendix F). Class-based instruction predominated, and this instruction was tested exclusively with older students (2nd through 6th graders) many of whom were poor readers. These conditions may have made it harder to produce substantial growth in reading.

Although there appear to be sizeable differences in effect sizes distinguishing the programs, the statistical test was not significant. However, drawing the conclusion that these programs are equally effective is premature because there were too few comparisons assessing each program to yield reliable results. Rather, findings should be considered suggestive in need of more studies for verification.

Evaluation of these separate programs was undertaken in the meta-analysis solely because of their prevalence in the database. The programs are older and hence more frequently studied than newer programs. But this does not mean that they are considered to be any better than newer programs that were not analyzed.

Impact of Synthetic Phonics Programs on Different Groups of Readers

Because there were so many comparisons (39) assessing the effects of synthetic phonics programs, it was possible to examine whether this type of program was more beneficial for some grade and reader ability groups than for others. Two groups, at-risk kindergartners and at-risk 1st graders, had the same effect size so they were combined into one group comprising nine comparisons. As evident in Table 3, all groups but one showed effect sizes significantly greater than zero, and all but one group had homogeneous sets of effects. This indicates that synthetic phonics programs produced stronger growth in reading than control programs in most of the different reader groups. Possible reasons why low-achieving readers in 2nd through 6th grades did not benefit were suggested earlier.

Effect sizes varied across the groups. A test to determine whether some groups benefited more from synthetic phonics than other groups showed that effects were significantly greater for at-risk kindergartners and first graders ($\underline{d} = 0.65$) than for the two groups of older 2nd through 6th grade readers. These findings indicates that synthetic phonics programs were especially effective for younger, at-risk readers.

Instructional Delivery Unit

Another property of systematic phonics instruction expected to influence growth in reading was the delivery unit. Three types were distinguished. There were eight treatments in which students received one-to-one tutoring. This was expected to be the most effective form of phonics instruction, particularly for low achieving and disabled readers, because it was tailored to individual students. Small group instruction was also expected to be especially effective because attention to individual students was still possible, and in addition, the social setting was expected to enhance motivation to perform and opportunities for observational learning. In the Panel's review of

phonemic awareness training studies, findings indicated that effect sizes were significantly greater with small groups than with classrooms or tutoring. Because classrooms involve a much higher ratio of students to teachers, phonics instruction delivered in this setting was expected to be less effective than in the other two settings.

In categorizing studies, it was easiest to determine when tutoring was used because this was clearly stated and described. Identifying whether studies used small groups was also straightforward because training procedures included this descriptive although it was not always clear that this was the only way that instruction was delivered. However, in the case of whole class instructional, sometimes this category was attributed to studies by default. In many reports, descriptions made clear that the phonics program was taught by teachers to their classrooms of students, but the unit of instruction they used to teach the phonics part of programs was not explicitly stated; so, it was inferred to be the class.

Before the meta-analysis was conducted, an adjustment was made to one effect size in the tutoring comparisons. This was considered important because there were only eight comparisons in this set. One of the tutoring studies (Tunmer & Hoover, 1993) produced an atypical effect size, $\underline{d} = 3.71$, which was much larger than the other effects. To limit the influence of this comparison on the overall mean, its value was reduced to equal the next largest effect size in the set, $\underline{d} = 1.99$.

Results of the analysis of effect sizes for the three types of instructional units revealed that all produced positive effects that were statistically greater than zero, indicating that tutoring, small groups and classes were all effective ways to deliver phonics instruction to students (see Table 3). In addition, the set of effect sizes for comparisons involving small groups was homogeneous, indicating that small group effects are not explained by additional moderator variables and that the mean is a good estimate of the actual effect size, $\underline{d} = 0.43$.

Tutoring produced an effect size of $\underline{d} = 0.57$ which was greater than the effect size for small groups, $\underline{d} = 0.43$, and for classrooms, $\underline{d} = 0.39$. However, none of these effects differed statistically from each other. This evidence falls short in supporting the expectation that

tutoring would prove especially effective for teaching phonics. However, perhaps there were too few comparisons assessing the effects of tutoring (only eight) to yield reliable findings. On the other hand, it might be noted that the instructional delivery given to the control groups against which tutoring was compared did not involve tutoring in the majority (62%) of the cases. This inequality should have given tutoring an extra advantage. However, it did not.

Inspection of effect sizes for individual studies in Table 2 reveals that some whole class programs produced effect sizes as large, and sometimes larger, than those produced by small groups or tutoring. Given the enormous expense and impracticality of delivering instruction in small groups or individually—except for children who have serious reading difficulties—research is needed to determine what makes whole class phonics instruction effective.

It is interesting to note that the same comparison of instructional units was conducted in the meta-analysis of phonemic awareness training effects. Results showed that small groups were significantly more effective than tutoring or classrooms. Why small groups were more effective for teaching phonemic awareness but not phonics is not clear and awaits further research.

Type of Control Group

To test whether systematic phonics programs produced superior growth in reading, researchers utilized control groups that received unsystematic phonics or no-phonics instruction. The types of control groups chosen by researchers varied across the studies. As mentioned earlier, some studies included more than one type of control group. Selected for analysis were the control groups that were taught the least amount of phonics. These were categorized into five types based on descriptions and labels provided in the studies: basal, regular curriculum, whole language, whole word, and miscellaneous.

Usually basal programs were those already in use at schools. "Regular curriculum" was the label covering cases in which controls received the traditional curriculum or the regular class curriculum in use at schools with no further specification of its contents other that asserting it did not teach phonics systematically. This category covered cases where performance in that grade at that school during previous

years was used as a baseline without additional description of the actual program taught. In comparisons involving students identified as at risk by schools, control groups received the standard intervention offered by the schools to treat reading problems.

Whole language was the label used by authors to characterize programs. In two studies (Freppon, 1991; Klesius et al., 1991), the purpose was to examine the effectiveness of whole language programs, not phonics programs that were taught to control groups. In both cases, phonics was taught with a "skill and drill" basal program that was not well described. Control groups that were taught with a Big Books program and with language experience were labeled as whole language.

There were a few programs given to control groups that taught whole words or sight words without much attention to letter-sound relations. These were classed as whole word programs.

Control group programs that did not fit into one of these categories were placed in a miscellaneous category. These included programs teaching traditional spelling, academic study skills, and tutoring in academic subjects. In one case, as a control for parents teaching their own children systematic phonics, the children spent time reading books to their parents (Leach & Siddall, 1990).

Of interest was whether phonics instruction would produce superior growth in reading regardless of the type of control group, and whether phonics instruction would appear more effective when compared to some types of control groups than to others. There were no a priori reasons for expecting effect sizes to be influenced by the type of control group, particularly since the criteria of standard-classroom instruction with minimal phonics had been applied consistently across studies in selecting control groups.

Results in Table 3 (Appendix E) reveal that all of the control groups yielded effect sizes that were statistically greater than zero and all favored the phonics treatment. Effects sizes ranged from $\underline{d} = 0.31$ for whole language controls to $\underline{d} = 0.51$ for whole word controls. Effect sizes for basal and miscellaneous control groups were homogeneous. Additional tests revealed that none of the

effect sizes differed significantly from the others. These findings indicate that systematic phonics instruction proved effective regardless of the type of control group that was used.

Design of Studies

Studies in the database varied in methodological rigor. It is important to rule out the possibility that the positive effects of phonics instruction detected in the meta-analysis arose from poorly designed studies. Three features of the studies were coded and analyzed to determine whether more rigorous designs yielded larger or smaller effect sizes: assignment of participants to treatment and control groups, potential presence of pre-experimental differences between groups, and sample size.

Random Assignment

Experimental designs that randomly assign students to treatment and control groups have stronger internal validity than designs that assign already existing groups to the treatment and control conditions. The latter procedure is referred to as nonequivalent group assignment. The goal of experiments is to provide solid evidence that the treatment or lack of it, rather than anything else, explains gains observed in performance following the treatment. Random assignment serves to reduce the likelihood that pre-experimental differences, rather than treatment effects, explain differences between treatment and control groups on outcome measures. When nonequivalent groups are used, statistical techniques can be applied to eliminate pretest differences between groups when outcome measures are analyzed. However, this is not as satisfactory a solution as random assignment.

Most studies in the database provided information regarding how students were assigned to treatment and control groups. If this was not mentioned, then the study was considered to have used nonequivalent groups. Table 3 (Appendix E) shows that studies using random assignment and studies using nonequivalent groups yielded very similar effect sizes, both of which were statistically greater than zero. These findings confirm that the positive effects of systematic phonics instruction did not arise primarily from studies with weaker nonequivalent group designs.

Pre-Experimental Differences

Studies were also coded for the presence of possible or actual pretest differences between treatment and control groups. Effect sizes for questionable studies were calculated separately from studies that were not questionable in this regard. There were 15 comparisons for which no information about pretests was provided and the groups were not randomly assigned. The mean effect size was $\underline{d} = 0.49$. There were ten studies that reported pretest differences and did not use random assignment. The mean effect size in this case was $\underline{d} = 0.37$. When studies containing potential or actual pretest differences were removed from the dataset, effect sizes changed very little and in fact increased slightly, from $\underline{d} = 0.44$ to $\underline{d} = 0.46$. These findings indicate that pretreatment differences between experimental and control groups did not explain why phonics-trained groups outperformed control groups on outcome measures across studies. It was the phonics instruction itself that very likely produced the greater gains in reading.

Sample Size

Another factor indexing the rigor of studies and the reliability of outcomes is sample size, with results of larger studies producing stronger results than smaller studies. The number of students participating in comparisons included in the database varied from 20 to 320. Sample sizes were used to group the comparisons into quartiles, and effect sizes were calculated for each quartile. From Table 3, it is apparent that effect sizes were very similar across quartiles and were all statistically greater than zero. The largest effect size, $\underline{d} = 0.49$, emerged in studies having the largest samples. These findings show that the positive effects of systematic phonics instruction were not limited to studies that produced effects with relatively few students.

Discussion

Findings of the meta-analysis allow us to conclude that systematic phonics instruction produces gains in reading and spelling not only in the early grades (kindergarten and 1st grades) but also in the later grades (2nd through 6th grades) and among children having difficulty learning to read. Effect sizes in the early grades were

significantly larger ($d = 0.55$) than effect sizes above 1st grade ($d = 0.27$). These results support Chall's (1967) assertion that early instruction in systematic phonics is especially beneficial to growth in reading.

Although there was some thought that kindergartners might not be ready for phonics instruction because they first need to acquire extensive knowledge about how print works (e.g., Stahl & Miller, 1989; Chall, 1996a, b), findings did not support this possibility. Phonics instruction produced similar effect sizes in kindergarten ($d = 0.58$) and 1st grade ($d = 0.54$).

Phonics instruction can be described in terms of the method used to teach children about letter-sound relations and how to use letter-sounds to read or spell. There are synthetic, analytic, analogy, spelling-based, and embedded approaches to teaching phonics. Phonics instruction can also be described in terms of the content covered, for example, short vowels, long vowels, digraphs, phonics generalizations, onsets and rimes, phonograms, and so forth. In the present meta-analysis, only the types of methods were compared in terms of the effect sizes produced, and no significant differences among methods were detected.

Stahl et al. (1998) suggest that the benefits of phonics instruction and differences among phonics approaches may arise from the amount of content covered and learned by students rather than from properties distinguishing the various methods. Synthetic methods tend to be efficient in covering content and tend to cover an ambitious number of sound-symbol correspondences in the 1st grade year. Other approaches vary considerably in the amount that they cover. To understand phonics instruction and its effects on student learning, research is needed to study separately the effects of teaching methods from the effects of content coverage. Systematic phonics instruction is focused on teaching children the alphabetic system and explicitly how to apply it to read and spell words. Phonics skills would be expected to show effects on text comprehension to the extent that phonics skills help children read the words in texts. This is one reason why phonics instruction may have exerted less impact on text comprehension outcomes than on word reading outcomes, because the impact is indirect. In addition, although phonics programs do give children practice reading connected text, the purpose of this practice is centered on word recognition rather than on

comprehending and thinking about the meaning of what is being read. This may be another reason why effect sizes on text comprehension were smaller than effect sizes on word reading.

In the present analysis, systematic phonics instruction exerted a lower than expected impact on reading growth in low achieving readers ($d = 0.15$) and disabled readers ($d = 0.32$). The Panel's meta-analysis of phonemic awareness training studies included comparisons involving poor readers. Most of these studies would qualify as phonics studies because letter-sound manipulations were part of the phonemic awareness training. The studies were not included in the phonics database in order to avoid duplication of studies across meta-analyses. The effect size on reading outcomes in the PA meta-analysis involving poor readers was $d = 0.45$, a value quite a bit higher than the effect sizes produced by phonics instruction. It may be that including more phonemic awareness training with letters might improve the quality of phonics instruction given to poor readers. However, there may be other factors that explain the difference as well. Closer scrutiny of the two sets of studies is needed to identify possible reasons. For example, RD students in the phonics analysis may have been older than students in the PA analysis.

The overall effect size of systematic phonics instruction in 1st grade was $d = 0.54$. Although moderate in size, this value is somewhat low when compared to effect sizes found in other similar reviews. Stahl and Miller (1989) conducted a meta-analysis of phonics instruction and drew their comparisons from the Cooperative First Grade Studies (Bond & Dykstra, 1967, 1998) whose participants should be similar to 1st graders in the present database. Stahl and Miller found effect sizes of 0.91 on the Stanford Word Reading subtest and 0.36 on the Paragraph Meaning subtest for children who received phonics instruction similar to that studied here. Overall, these are higher effect sizes than those detected in the present meta-analysis.

The discrepancy may arise from differences in the way the Panel created its database. Whereas the Panel's review was limited to studies published in peer-reviewed journals, authors of the previous meta-analyses made a great effort to find "fugitive" or unpublished studies to include. One reason to search widely for studies is that the publishing process tends to

screen out studies reporting null effects, and this runs the risk of biasing the data set towards positive effects. However, such a bias would be expected to favor a larger effect size using National Reading Panel procedures, and this did not happen. Another possible reason for the discrepancy is that the previous analyses included unpublished studies, thus running the risk of admitting studies of poor quality with inflated effect sizes. Limiting studies to those passing the test of peer review minimizes this risk.

Another possible explanation for the Panel's smaller effect size is that the database involved more recent studies. There may have been more of a tendency for later studies to focus on at-risk, low-achieving, and disabled readers for whom growth in reading may be harder to achieve. Perhaps the reading instruction experienced by students in control groups included more phonics than the reading instruction received by control groups in earlier years. In the 1960s, basal readers used a whole word methodology whereas the control conditions in more recent studies are presumably more eclectic. Table 2 identifies the control groups used by studies in the corpus. Whereas some groups were true "no-phonics" controls, other groups received some phonics instruction. It may be that, instead of examining the difference between phonics instruction and no phonics instruction, a substantial number of studies actually compared more systematic phonics instruction to less phonics instruction. This would produce smaller differences between treatment and control groups and hence smaller effect sizes.

In one of the studies in the database, Evans and Carr (1985) conducted extensive observations of the instruction received by treatment and control groups and reported their observations numerically. They found that the phonics classes spent 13.38% of the group time and 11.94% of independent work time on word analysis, whereas the control group spent 5.37% of the group time and 1.84% of the independent time on word analysis. Although there is a difference favoring the phonics group, the finding shows that control classes did spend some time on word analysis as well. Chall and Feldmann (1966) found that there was considerable variation in instruction, even in classes professing to be using the same methods. This underscores the importance of researchers taking steps not only to

assess outcomes of instructional treatments but also to document the nature of the instruction received by treatment and control groups to verify whether and how they actually differed.

Studies to Illustrate Systematic Phonics Instruction and Its Contribution to Growth in Reading

Some of the studies in the database are described to provide a glimpse of the experiments contributing effect sizes and to portray various types of phonics instruction that were examined.

Phonics Instruction in Kindergarten

Systematic phonics instruction in kindergarten was studied in six articles. The main goals included teaching children the shapes of letters and their sounds, how to analyze sounds in words (phonemic awareness), and how to use letter sounds to perform various reading or writing tasks appropriate for children just starting out. In the study by Stuart (1999), three kindergarten teachers utilized the Jolly Phonics program (Lloyd, 1993), and three teachers centered their instruction around Holdaway's (1979) Big Book approach. Teachers taught these programs 1 hour per day for 12 weeks during the latter half of kindergarten.

Big Book instruction included work with letters. Teachers drew children's attention to written words in the books and they talked about letters in words. Also, teachers employed various "imaginative and fun activities" to help children learn letters and their sounds. However, the instruction was not systematic; the sequence of teaching letters was not prescribed, and no special system for remembering letter-sound relations was taught.

The Jolly Phonics program was more systematic and prescribed in its teaching of letters. This program was developed by Lloyd (1993), a teacher, for 4- and 5-year-olds in their first year of schooling in the United Kingdom. Central to the program is the use of meaningful stories, pictures, and actions to reinforce recognition and recall of letter-sound relationships, and precise articulation of phonemes. There are five key elements to the program: (1) learning the letter sounds, (2) learning letter formation, (3) blending for reading, (4) identifying the sounds in words for writing, and (5) tricky words that are high frequency and irregularly

spelled. The program includes activities and instruction specifically designed to address those skills most needed in the development of early literacy. Unlike many older phonics approaches, however, Jolly Phonics promotes playful, creative, flexible teaching that fits well with whole language practice and leads directly to authentic reading and writing.

At the end of training in either Jolly Phonics or Big Books, children were given various tests to compare effects of the programs. Results showed that Jolly Phonics at-risk kindergartners were able to read significantly more words and pseudowords and to write more words than the Big Book group. The overall effect size was $\underline{d} = 0.73$. A year later, the children were retested. The Jolly Phonics group outperformed the control group in reading and spelling words but not in reading comprehension. These results show that phonics instruction in kindergarten is effective in boosting children's progress in learning to read and write words.

One interesting feature of the Jolly Phonics program is that children are taught hand gestures to help them remember the letter-sound associations. For example, they make their fingers crawl up their arm portraying an ant as they chant the initial sound of "ant" associated with the letter $\underline{a}$. The value of mnemonics for teaching letter-sound relations to kindergartners is supported by evidence. In a study by Ehri, Deffner, and Wilce (1984), children were shown letters drawn to assume the shape of a familiar object, for example, $\underline{s}$ drawn as a snake, $\underline{h}$ drawn as a house (with a chimney). Memory for the letter-sound relations was mediated by the name of the object. Children were taught to look at the letter, be reminded of the object, say its name, and isolate the first sound of the name to identify the sound (i.e., s - $\underline{snake}$ - /s/). With practice they were able to look at the letters and promptly say their sounds. Children who were taught letters in this way learned them better than children who were taught letters by rehearsing the relations with pictures unrelated to the letter shapes (e.g., house drawn with a flat roof and no chimney) and also better than children who simply rehearsed the associations without any pictures.

Application of this principle can be found in Letterland (Wendon, 1992), a program that teaches kindergartners letter-sound associations. In this program, all the letters are animate characters that assume the shape of the

letters and have names prompting the relevant sound, for example, Sammy Snake, Hairy Hat Man, Fireman Fred, Annie Apple. The task of learning the shapes and sounds of all the alphabet letters is difficult and time-consuming, particularly for children who come to school knowing none. The relations are arbitrary and meaningless. Techniques to speed up the learning process are valuable in helping kindergartners prepare for formal reading instruction.

The motivational value of associating letters with interesting characters or hand motions and incorporating this into activities and games that are fun is important for promoting young children's learning. If the task of teaching letters is stripped bare to one of memorizing letter shapes and sounds, children will become bored and easily distracted and will take much longer to learn the associations.

A Developmental Approach to Phonics Instruction in Kindergarten

Another phonics program for kindergartners was studied by Vandervelden and Siegel (1997). The interesting feature of their approach was to tailor the intervention to individual children's level of knowledge. This is important because kindergartners vary greatly in how much they already know about letters when they enter school. The instruction lasted 12 weeks, with children receiving two sessions per week. There were 15 children that received phonics instruction and 15 that received the same instructional format but focused on classroom activities and materials. Children were pretested. The three children who showed the least knowledge received one-on-one tutoring, the next eight lowest scoring children were instructed in pairs, and the four highest scoring children worked in a small group.

The skills taught to phonics-treated children who lacked them included the following: learning sounds for consonant letters; use of initial letter-sound matches to recognize, spell, and read words; segmenting words into sounds and spelling the sounds; orally reading text containing the words learned in this way; learning correct spellings of words by analyzing letter-sound constituents; and use of rime analogy in reading and spelling words. Easier skills were taught before harder skills. Instruction began at levels appropriate for individual learners.

In the control group, children engaged in activities used in their classrooms. This included letter learning and phonemic awareness. However, children were not explicitly guided in the use of these skills to read and write.

Results showed that the phonics groups outperformed the control group on tests of phonemic awareness and letter-sound relations but not letter names. Also, the phonics group did better on tests of speech-print matching of words and pseudowords (e.g., which written word, <u>milk</u>, <u>monk</u>, or <u>mask</u> says "mask"), on tests of writing the sounds in words, and on some but not all measures of word reading. The overall effect size was $\underline{d} = 0.47$. It is important to recognize, however, that these kindergartners were still at a rudimentary level in their development as readers. For example, at the end of the treatment, they were able to match 43% of the written and spoken words correctly; they read only a mean of 10 out of 60 high frequency words such as <u>up</u>, <u>yes</u>, and <u>book</u>, and they spelled only 46% of the sounds in words. This suggests that teaching students to use phonics skills to read and spell words at the kindergarten level may yield only limited success. However, perhaps this program was not optimally designed or did not last long enough.

A 2.5-Year Phonics Program Beginning With Phonemic Awareness

A lengthier, more comprehensive program lasting more than 2 years was studied by Blachman et al. (1999). Classroom teachers used the program with low SES, inner-city children. Instruction began in kindergarten with a focus on phonemic awareness training lasting 11 weeks. In 1st grade, explicit, systematic instruction in the alphabetic code was taught. This instruction continued in 2nd grade for children who did not complete the program in 1st grade. Control children participated in the school's regular basal reading program that included a phonics workbook that children used independently.

The phonemic awareness instruction taught children to perform a "say it and move it" procedure in which they moved a disk down a page as they pronounced each phoneme in a word. They practiced segmenting two- and three-phoneme words in this way. Then a limited set of eight letter-sound relations was taught, and children moved the letters rather than the disks. It is noteworthy that when children began this program, they

knew on average only two letter sounds and could not yet write their names. Thus, the participants were starting from zero in their alphabetic learning. By the end of kindergarten, children knew on average 19 letter names and 13 letter-sounds, indicating that substantial learning had occurred.

At the beginning of 1st grade, there was still wide variation in children's letter knowledge and phonemic awareness. This underscores the fact that even though children receive the same instruction, they still differ in how quickly they learn what they are taught. To address the variation, children were assigned to ability groups. The core of the reading program involved daily, 30-minute lessons consisting of five steps that emphasized the alphabetic code:

1. Teaching new sound-symbol correspondences with vowels highlighted in red

2. Teaching phoneme analysis and blending

3. Reading regularly spelled, irregularly spelled, and high-frequency words on flash cards to develop automaticity

4. Reading text containing phonetically controlled words

5. Writing four to six words and a sentence to dictation.

By the end of the program, children had been introduced to all six syllable types: closed (<u>fat</u>), final E (<u>cake</u>), open (<u>me</u>), vowel team (<u>pain</u>), vowel + r (<u>burn</u>), and consonant <u>le</u> (<u>table</u>). Vocabulary development and work on reading comprehension was incorporated as well, with more time spent reading text as the year progressed and children's reading vocabulary grew.

Inservice workshops held once a month were used to instruct teachers how to implement the program. The instruction presented information about how children acquire literacy skills and the role of phonological processes in learning to read. Teachers learned how to provide explicit instruction in the alphabetic code. The issue of pacing was stressed. Developing students' phonemic awareness, letter-sound knowledge, and word recognition skills was identified as being more important than "covering the material."

To assess how far children had progressed in their reading and writing, various tests were given at the end of kindergarten, 1st grade, and 2nd grade. Results showed that kindergartners receiving PA training outperformed control students, with $\underline{d} = 0.72$. At the end of 1st grade, children who received explicit phonics training achieved significantly higher scores than controls, with $\underline{d} = 0.64$. During 2nd grade, children in the phonics group who had not met the program's goals received additional instruction while the rest received regular classroom instruction. On posttests at the end of the year, the phonics-trained group continued to outperform the control group, with $\underline{d} = 0.36$.

These findings show that the explicit systematic instruction in phonics provided by the Blachman program improved low SES children's ability to read words more than a basal program less focused on teaching children alphabetic knowledge and word reading skills. Several features of this program are noteworthy and may underlie its effectiveness. The same program continued over three grades, thus insuring consistency and continuity in children's learning the alphabetic system and how to use it to read and spell. The program began in kindergarten with alphabetic code instruction that was appropriate for children's level of knowledge. They were taught phonemic awareness and a limited set of letter-sound relations which they used to make and break words. Both PA and letter knowledge are known to be the strongest predictors of how well children will succeed in learning to read. Delivery of instruction was tailored to enable all students to complete the program. Tests were given to assess children's progress and to distinguish those children who needed further instruction from those who did not. Instruction in the alphabetic code included various kinds of reading and writing skills, not only sounding out and blending words but also building memory for words, spelling words, and reading words in text. An extensive set of letter-sound relations including vowels was taught and applied to various types of words organized by syllable structure. Teachers were provided with inservice workshops during the school year to help them not only provide instruction correctly but also to understand the reading processes and their course of acquisition in students. These properties of the Blachman phonics program may account for its effectiveness. Further research to examine the contribution of such properties is needed.

An Intensive 3-Year Tutoring Program: Synthetic vs. Embedded Phonics Instruction

Another study in the database, by Torgesen et al. (1999), also provided phonics instruction throughout the primary grades. In this study, two different forms of phonics instruction were compared, one which provided very explicit and intensive instruction in PA and phonemic decoding called PASP (phonological awareness plus synthetic phonics), while the other provided systematic but less explicit instruction in phonemic decoding in the context of more instruction and practice in text comprehension, called EP (embedded phonics). Instruction was provided by tutors rather than classroom teachers. Kindergarten children with poor PA and letter knowledge received 88 hours of tutoring over 2.5 years, with sessions lasting 20 minutes and scheduled four times per week. Instruction was individually paced according to the progress that children made. This instruction was added to the reading instruction they received in the classroom. There were two control groups, one that received tutoring that supported regular classroom instruction, and one in which children received only regular classroom instruction. Instruction in the tutoring control condition included some phonics oriented activities. There were 180 children from 13 schools. Children were randomly assigned to one of the four conditions.

The PASP children received the Auditory Discrimination in Depth program (Lindamood & Lindamood, 1984). This program began by teaching children phonemic awareness in a unique way. Children were led to discover and label the articulatory gestures associated with each phoneme by analyzing their own mouth movements as they produced speech. For example, children learned that the word <u>beat</u> consists of a lip popper, a smile sound, and a tongue tapper. Children learned to track the sounds in words with mouth pictures as well as colored blocks and letters. Most of the time in this program was spent building children's PA and their decoding skills although some attention was given to the recognition of high frequency words, text reading, and comprehension.

The EP program began by teaching children to recognize whole words. Instruction in letter-sounds occurred in the context of learning to read words from memory (by sight). Also, children wrote sentences and read what they wrote. In this context, phonemic

awareness was taught by having children segment the sounds in words before writing them. When children had sufficient reading vocabulary, they began reading short stories to build their reading vocabulary further. The emphasis was on acquiring word level reading skills, including sight words and phonemic decoding skills. Also, attention was given to constructing the meanings of stories that were read.

One step taken in the Torgesen et al. study was to videotape 25% of the PASP and EP tutorial sessions and analyze the interaction to verify how phonics instruction differed in the two programs. The percentages of time spent on the following types of activity were

• PA, letter-sounds, phonemic reading/writing of words: 74% (PASP) vs. 26% (EP)

• Sight word instruction: 6% (PASP) vs. 17% (EP)

• Reading/writing connected text: 20% (PASP) vs. 57% (EP).

In comparing the groups' performance on outcomes measures across the grades, Torgesen et al. found that the PASP group read significantly more real words and nonwords and spelled more words than one or both of the control groups. However, the EP group did not outperform the control groups on any of the measures. There was a significant overall effect of interventions on the comprehension measures, but individual contrasts between groups were not statistically significant. Comparison of the PASP and EP groups revealed superior performance by PASP on measures of phonological awareness, phonemic decoding accuracy and efficiency, and word reading accuracy. However, the groups did not differ in word reading efficiency (taking account of speed as well as accuracy) or in the individual contrasts for reading comprehension. Thus, findings revealed that intensive training in phonics produced superior word reading skills compared to embedded phonics training or training given to control groups. Interestingly, neither of the two instructional control groups, embedded phonics or supported classroom instruction, produced significant effects

compared to the no treatment control group, while the explicit PASP group did. Based on comparisons to the classroom control group, effect sizes for the two phonics groups were

• PASP: $\underline{d}$ = 0.33 (kindergarten), 0.75 (1st grade), 0.67 (2nd grade)

• EP: $\underline{d}$ = 0.32 (kindergarten), 0.28 (1st grade), 0.17 (2nd grade).

Clearly, effects of synthetic phonics instruction persisted more strongly over the grades than effects of embedded phonics instruction. Left unclear is whether PASP's effectiveness resulted from the greater time spent teaching alphabetic and phonological processes, or the specific content of the instruction, or a combination of both factors.

Although the comparisons between individual groups were not significant for the comprehension measures, when the outcomes for the PASP group were compared to those of the EP and RCS groups combined, the effect size for the passage comprehension test of the Woodcock Reading Mastery Test-Revised was 0.43. The corresponding effect size for the comprehension measure for the Gray Oral Reading Test–3 was 0.21. While reading comprehension depends upon other processes besides word reading, one would expect to see transfer, particularly in the primary grades where text reading is heavily influenced by word recognition skills. One possible explanation is that the tests of comprehension were standardized and hence were not sufficiently sensitive to detect small within-grade differences. This is because standardized tests are designed to detect differences across the whole range of grades; so, there are only a small number of items at each grade level. Another possibility is that compensatory processes are sufficiently strong to dilute the contribution that superior word recognition skill makes to text reading. That is, children read and comprehend text by utilizing their linguistic and background knowledge combined with their word reading skill. When word reading skill is somewhat weaker, children can rely more heavily on their knowledge about the subject and memory for what they have read to still make sense of the text.

The kindergartners selected to be tutored in reading in the Torgesen et al. (1999) study were severely at risk for becoming disabled readers based on very poor letter knowledge and phonemic awareness which are the two best kindergarten-entry predictors of future reading achievement (Share et al., 1984). However, these children varied greatly in verbal intelligence, with IQs ranging from 76 to 126 in kindergarten and from 57 to 130 in 2nd grade. Thus, the sample in this study included two kinds of potentially poor readers, children who were *unexpectedly* poor readers because their IQs were *higher* than their reading potential scores and children whose below-average reading was not surprising given that their IQs were also below average. These two types of poor readers have been distinguished in other studies by researchers. Various labels, such as dyslexic or learning disabled or reading disabled, have been applied to children whose higher IQs are discrepant with their lower reading skill. Children whose lower reading scores are consistent with their lower IQs have been called low achievers or garden variety poor readers. These children would be expected to display low achievement not only in reading but also in other academic areas requiring cognitive or verbal capabilities.

Torgesen et al. (1999) observed that children in their study varied greatly in their response to instruction. Even in the strongest phonics group, almost one-fourth of the children remained significantly impaired in their decoding or word reading ability at the end of instruction. Torgesen et al. conducted a regression analysis to examine what characteristics of the children predicted how well or poorly they responded to instruction as indexed by their growth in word reading over 2.5 years. They found that the important variables explaining growth were home background (parent occupation and education), kindergarten classroom behavior (activity level, attention, adaptability, social behavior) and phonological capabilities (i.e., phonemic awareness, short-term memory, naming speed). The variable involving IQ differences among the children did not explain any further growth over and above these other variables. Torgesen et al. suggest that whether or not children's IQ is discrepant with their reading potential is probably not relevant in determining their need for special help in acquiring word reading skills.

Modified Reading Recovery© Studies

There were three studies in the database that adopted the Reading Recovery© (RR) format developed by Clay (1993) and altered it to include more systematic work in phonics. The type of phonics instruction involved an emphasis on larger subunits as well as phonemes. The RR program developed by Clay is adminstered by a tutor to children who have fallen behind in reading after a year of instruction. The 30-minute RR lesson includes several activities: rereading two familiar books, reading the previous day's new book, practicing letter identification, writing a story by analyzing sounds in words, re-assembling the words of a cut-up story, reading a new book.

Greaney, Tunmer, and Chapman (1997) modified the RR program by providing explicit instruction in letter-phoneme patterns once children had learned the majority of letters. This work consumed 5 minutes of each session and was substituted for the letter segment of the RR lesson. Children were taught to read pairs of nouns containing common spellings of rimes (e.g., m-eat) and then words with the rime embedded in it (e.g., h-eat-er). They practiced reading and also writing words with these larger rime units referred to as "eggs" because the unit was written in an egg-shaped space. Attention was drawn to the egg units and their utility for reading words. During the final book reading segment of each session, children were encouraged to use the eggs to identify unfamiliar words in the book. This treatment was referred to as rime analogy training. Children in the control group followed the same RR format and read the same words. However, no attention was drawn to rime units in the words, and the words were mixed up rather than taught in sets having the same rimes.

The study was conducted in New Zealand. Both the modified RR and the unmodified RR programs lasted for 12 weeks. The children in the study were from grades 2 through 5 and were the poorest readers in their class. Results showed that the children who received rime training outperformed control children on tests of word and pseudoword reading but not on tests of reading comprehension. The overall effect size was $\underline{d} = 0.37$. These findings reveal that the rime-analogy phonics program produced greater growth in word reading than the whole word program.

Tunmer and Hoover (1993) performed a similar study in which the letter segment of the RR lesson was replaced by more systematic phonics instruction. Children were taught to make, break, and build new words that had similar letters and sounds. Instruction began by focusing on phonograms or rime spellings in words (e.g., <u>make</u>, <u>bake</u>, <u>cake</u>, <u>take</u>). A metacognitive strategy training approach rather than a skill and drill approach was used to make children aware of how letters and sounds work in words and how to use their alphabetic knowledge to read and spell.

Two control groups were included in the study. One group received unmodified RR lessons. The other group received the standard treatment given to poor readers by the school district. This was a pull-out program in which teachers worked with children in small groups. Some word analysis activities were included. The children were all 1st graders in their 2nd year of reading instruction. They were the poorest readers in their class. Posttests were given when RR children achieved the goals of the program. Results showed that the modified RR group outperformed the group receiving the standard small-group instruction on all measures. The overall effect size was $\underline{d} = 3.71$, indicating that the modified RR phonics program produced an enormous advantage over the treatment received by the standard control group.

In contrast, the modified RR group performed very similarly to the unmodified RR group on the reading measures following training. The only difference was that it took significantly fewer sessions for the modified RR group to achieve the goals of RR than the unmodified RR group. The effect size showing the advantage in reduced time was $\underline{d} = 1.40$. The same advantage in time, but not in reading outcomes, was uncovered by Iversen and Tunmer (1993) who conducted a very similar study. (The Iversen and Tunmer data were included in the Panel's meta-analysis of phonemic awareness instruction.) Findings of both studies show that Clay's Reading Recovery© program produced the same growth in reading even though it provided less systematic phonics instruction than the modified program and provided it mainly through writing exercises rather than decoding activities. Although reading outcomes were the same, the fact that one program took less time makes the more intensive phonics approach preferable. Because the RR program requires one-on-one tutoring delivered in schools by a few highly trained RR teachers, it is expensive; so, a savings in time can mean either that more students are helped or that fewer teachers are required.

A third study in the database also modified the RR format to include more systematic phonics instruction. In the study by Santa and Hoien (1999), at-risk 1st graders received tutoring that involved story reading, writing, and phonological skills based on a program developed by Morris (1992). The unique part of this phonics program was that it used word study activities to develop phonological awareness and decoding skill. Word study consumed 5 to 6 minutes of the 30-minute lesson. Children were given cards to sort into categories. They might sort picture cards that shared the same initial sounds, or word cards sharing the same vowel sounds. The typical sort involved three patterns with four words in each pattern. Initially, children worked with phonograms (e.g., -<u>at</u> in <u>hat</u>, <u>cat</u>, <u>sat</u>, <u>rat</u>) and then advanced to shared phonemes as the basis for sorting words. Children also were taught to spell by writing letters for the sounds heard in words. Metacognitive strategies were taught including an analogy strategy in which children were urged to use words they know to read words they don't know.

The control group received small group, guided reading instruction. They practiced reading and rereading books in 30-minute lessons but did not receive any word study activities. Results showed that the word study program produced much greater growth in reading than the guided reading program, $\underline{d} = 0.76$. Gains were greater in reading comprehension as well as word reading. These findings provide evidence for the effectiveness of teaching children phonics through the use of larger units along with phonemes.

Systematic Phonics to Remediate the Reading Difficulties of Disabled Readers

Children who have been diagnosed as reading disabled have severe reading difficulties that are not explained by low intelligence. Systematic phonics programs have been developed to remediate their reading difficulties. RD children have special problems in acquiring word reading skills. Not only do they struggle to read pseudowords, but they also have trouble remembering how to read words they have read before.

Maureen Lovett and her associates (Lovett et al., 1994; Lovett & Steinbach, 1997; Lovett et al., in press) have conducted several studies to examine how to improve the word reading skills of severely disabled readers. They have explored the effectiveness of two types of phonics programs, a synthetic program they call PHAB and a larger-unit program, which teaches children to use subparts of words they know to read new words, referred to as WIST.

The PHAB synthetic phonics program adopted the Direct Instruction model developed by Engelmann and his colleagues (see Appendix) to remediate the decoding and phonemic awareness difficulties of the disabled readers. Children were taught to segment and blend words orally. They were taught letter-sound associations in the context of word recognition and decoding instruction. The program taught a left-to-right decoding strategy to sound out and blend letters into words. Special marks on letters and words provided visual cues to aid in decoding, such as symbols over long vowels, letter size variations, and connected letters to identify digraphs. Cumulative, systematic review and many opportunities for overlearning were used. New material was not introduced until the child had fully mastered previously instructed material. Children were taught in small groups.

The larger-unit, word analogy program called WIST was adapted from the Benchmark Word Identification/ Vocabulary Development program developed by Gaskins et al. (1986). This program had a strongly metacognitive focus. It taught children how to use four metacognitive strategies to decode words: reading words by analogy, detecting parts of words that are known, varying the pronunciations of vowels to maintain flexibility in decoding attempts, and "peeling off" prefixes and suffixes in words. Children learned a set of 120 key words exemplifying high-frequency spelling patterns, five words per day. They learned to segment the words into subunits so that they could use known words and their parts to read other similarly spelled words. They learned letter-sound associations for vowels and affixes. Various types of texts provided children with practice applying the strategies that were taught.

The children participating in the studies were referred to Lovett's clinic because they had severe reading problems. Children were randomly assigned to receive the PHAB program, the WIST program, or a non-reading control program that involved teaching students academic survival skills such as organization and problem solving relevant to the classroom. The students ranged in age from 6 to 13 years or grades 2nd through 6th. The three programs took the same amount of time. In one study, it was 35 hours; in another study, 70 hours.

To evaluate the effectiveness of the programs, performance of students receiving either PHAB or WIST were compared to performance of the control group. There were four comparisons assessing effects of PHAB and four assessing WIST in the database. Although the effect sizes were somewhat variable, the average effect size across the comparisons indicated that both programs produced about the same growth in reading, $\underline{d} = 0.41$ for PHAB and $\underline{d} = 0.48$ for WIST. In two of the comparisons, both reading comprehension and word reading were measured. Substantial gains were evident on both measures. These findings indicate that the two approaches to teaching systematic phonics, one teaching synthetic phonics, and one teaching the use of larger subunits of words to read by analogy, were quite effective in helping disabled readers improve their reading skills.

Conclusions

There were 38 studies from which 66 treatment-control group comparisons were derived. Although each comparison could contribute up to six effect sizes, one per outcome measure, few studies did. The majority (76%) of the effect sizes involved reading or spelling single words, whereas 24% involved text reading. The imbalance favoring single words is not surprising given that the focus of phonics instruction is on improving children's ability to read and spell words. Studies limiting instructional attention to children with reading problems accounted for 65% of the comparisons, 38% involving poor readers considered "at risk" or low achieving and 27% diagnosed as reading disabled (RD). Studies involving 1st graders were overrepresented in the database, accounting for 38% of the comparisons. Fewer kindergartners (12%) and children in 2nd through 6th grades (23%) were represented. Children in

the RD group spanned several ages and grades, ranging from ages 6 to 13 and grades 2nd through 6th. Most of the studies (72%) were recently conducted, in the past 10 years.

Systematic phonics instruction typically involves explicitly teaching students a prespecified set of letter-sound relations and having students read text that provides practice using these relations to decode words. Instruction lacking an emphasis on phonics instruction does not teach letter-sound relations systematically and selects text for children according to other principles. The latter form of instruction includes whole-word programs, whole language programs, and some basal reader programs.

The meta-analyses were conducted to answer several questions about the impact of systematic phonics instruction on growth in reading when compared with instruction that does not emphasize phonics. Findings provided strong evidence substantiating the impact of systematic phonics instruction on learning to read.

1. Does systematic phonics instruction help children learn to read more effectively than unsystematic phonics instruction or instruction teaching no phonics?

Children's reading was measured at the end of training if it lasted less than a year or at the end of the first school year of instruction. The mean overall effect size produced by phonics instruction was significant and moderate in size ($\underline{d} = 0.44$). Findings provided solid support for the conclusion that systematic phonics instruction makes a more significant contribution to children's growth in reading than do alternative programs providing unsystematic or no phonics instruction.

2. Are some types of phonics instruction more effective than others? Are some specific phonics programs more effective than others?

Three types of phonics programs were compared in the analysis: (1) synthetic phonics programs that emphasized teaching students to convert letters (graphemes) into sounds (phonemes) and then to blend the sounds to form recognizable words; (2) larger-unit phonics programs that emphasized the analysis and

blending of larger subparts of words (i.e., onsets, rimes, phonograms, spelling patterns) as well as phonemes; and (3) miscellaneous phonics programs that taught phonics systematically but did this in other ways not covered by the synthetic or larger-unit categories or were unclear about the nature of the approach. The analysis showed that effect sizes for the three categories of programs were all significantly greater than zero and did not differ statistically from each other. The effect size for synthetic programs was $\underline{d} = 0.45$; for larger-unit programs, $\underline{d} = 0.34$; and for miscellaneous programs, $\underline{d} = 0.27$. The conclusion supported by these findings is that various types of systematic phonics approaches are more effective than non-phonics approaches in promoting substantial growth in reading.

There were seven programs that were examined in three or more treatment-control group comparisons in the database. Analysis of the effect sizes produced by these programs revealed that all were statistically greater than zero and none differed statistically from the others in magnitude. Effect sizes ranged from $\underline{d} = 0.23$ to 0.68. In most cases there were only three or four comparisons contributing effect sizes, so results may be unreliable. The conclusion drawn is that specific systematic phonics programs are all more effective than non-phonics programs and they do not appear to differ significantly from each other in their effectiveness although more evidence is needed to verify the reliability of effect sizes for each program.

3. Is phonics taught more effectively when students are tutored individually, when they are taught in small groups, or when they are taught as classes?

All three delivery systems proved to be effective ways of teaching phonics, with effect sizes of $\underline{d} = 0.57$ (tutoring), $\underline{d} = 0.43$ (small group), and $\underline{d} = 0.39$ (whole class). All effect sizes were statistically greater than zero, and no one differed significantly from the others. This supports the conclusion that systematic phonics instruction is effective when delivered through tutoring, through small groups, and through teaching classes of students.

4. Is phonics instruction more effective when it is introduced to students not yet reading, in kindergarten or 1st grade, than when it is introduced in grades above 1st after students have already begun to read?

Phonics instruction taught early proved much more effective than phonics instruction introduced after 1st grade. Mean effect sizes were kindergarten $\underline{d}$ = 0.56; 1st grade $\underline{d}$ = 0.54; and 2nd through 6th grades $\underline{d}$ = 0.27. The conclusion drawn is that systematic phonics instruction produces the biggest impact on growth in reading when it begins in kindergarten or 1st grade before children have learned to read independently. To be effective, phonics instruction introduced in kindergarten must be appropriately designed for learners and must begin with foundational knowledge involving letters and phonemic awareness.

5. Is phonics instruction beneficial for children who are having difficulty learning to read? Is it effective in preventing reading failure among children who are at risk for developing reading problems in the future? Is it effective in remediating reading difficulties in children who have been diagnosed as reading disabled and children who are low-achieving readers?

Phonics instruction produced substantial reading growth among younger children at risk of developing future reading problems. Effect sizes were $\underline{d}$ = 0.58 for kindergartners at risk and $\underline{d}$ = 0.74 for 1st graders at risk. Phonics instruction also improved the reading performance of disabled readers (i.e., children with average IQs but poor reading) for whom the effect size was $\underline{d}$ = 0.32. These effect sizes were all statistically greater than zero. However, phonics instruction failed to exert a significant impact on the reading performance of low-achieving readers in 2nd through 6th grades (i.e., children with reading difficulties and possibly other cognitive difficulties explaining their low achievement). The effect size was $\underline{d}$ = 0.15, which was not statistically greater than chance. Possible reasons might be that the phonics instruction provided to low-achieving readers was not sufficiently intense, that their reading difficulties arose from sources not treated by phonics

instruction such as poor comprehension, or that there were too few cases (i.e., only eight treatment-control comparisons pulled from three studies) to yield reliable findings.

The conclusion drawn from these findings is that systematic phonics instruction is significantly more effective than non-phonics instruction in helping to prevent reading difficulties among at-risk students and in helping to remediate reading difficulties in disabled readers. No conclusion is drawn in the case of low-achieving readers because it is unclear why systematic phonics instruction produced little growth in their reading and whether the finding is even reliable. Further research is needed to determine what constitutes adequate remedial instruction for low-achieving readers.

6. Does systematic phonics instruction improve children's reading comprehension ability as well as their decoding and word-reading skills?

Systematic phonics instruction was most effective in improving children's ability to decode regularly spelled words ($\underline{d}$ = 0.67) and pseudowords ($\underline{d}$ = 0.60). This was expected because the central focus of phonics programs is upon teaching children to apply the alphabetic system to read novel words. Phonics programs also produced growth in the ability to read irregularly spelled words although the effect size was significantly lower, $\underline{d}$ = 0.40. This is not surprising because a decoding strategy is less helpful for reading these words. However, alphabetic knowledge is useful for establishing connections in memory that help children read irregular words they have read before. This may explain the contribution of phonics.

Systematic phonics instruction produced significantly greater growth than non-phonics instruction in younger children's reading comprehension ability ($\underline{d}$ = 0.51). However, the effects of systematic phonics instruction on text comprehension in readers above 1st grade were mixed. Although gains were significant for the subgroup of disabled readers ($\underline{d}$ = 0.32), they were not significant for the older group in general ($\underline{d}$ = 0.12).

The conclusion drawn is that growth in word-reading skills is strongly enhanced by systematic phonics instruction when compared to non-phonics instruction for kindergartners and 1st graders as well as for older

struggling readers. Growth in reading comprehension is also boosted by systematic phonics instruction for younger students and reading disabled students. Whether growth in reading comprehension is produced generally in students above 1st grade is less clear.

7. Does systematic phonics instruction have an impact on children's growth in spelling?

Systematic phonics instruction produced much growth in spelling among the younger students, that is, kindergartners and 1st graders, $\underline{d} = 0.67$, but not among the older students above 1st grade, whose effect size of $\underline{d} = 0.09$ did not differ from zero. One factor contributing to the difference is that younger children were given credit for using phonics-based knowledge to produce letter-sound spellings of words as well as correct spellings whereas older children were not. Another factor may be that as children move up in the grades, remembering how to spell words requires knowledge of higher level regularities not covered in systematic phonics programs. A third reason for the poor showing among older students may be that the majority were poor readers who are known to have difficulty learning to spell.

The conclusion drawn is that systematic phonics instruction contributed more than non-phonics instruction in helping kindergartners and 1st graders apply their knowledge of the alphabetic system to spell words. However, it did not improve spelling in students above 1st grade.

8. Is systematic phonics instruction effective with children at different socioeconomic levels?

Systematic phonics instruction helped children at all SES levels make greater gains in reading than did non-phonics instruction. The effect size for low-SES students was $\underline{d} = 0.66$, and for middle-class students it was $\underline{d} = 0.44$. Both were statistically greater than zero and did not differ from each other. The conclusion drawn is that systematic phonics instruction is beneficial to students regardless of their socioeconomic status.

9. Does the type of control group used to evaluate the effectiveness of systematic phonics instruction make a difference?

The type of nonsystematic or non-phonics instruction given to control groups to evaluate the effectiveness of systematic phonics instruction varied across studies and included the following types: basal programs, regular curriculum, whole language approaches, whole word programs, and miscellaneous programs. The question of whether phonics produced better reading growth than each type of control group was answered affirmatively in each case. The effect sizes were all positive favoring systematic phonics, were all statistically greater than zero, and ranged from $\underline{d} = 0.31$ to 0.51. No single effect size differed from any of the others.

The conclusion supported by these findings is that the effectiveness of systematic phonics instruction found in the present meta-analysis did not depend on the type of instruction that students in the control groups received. Students taught systematic phonics outperformed students who were taught a variety of nonsystematic or non-phonics programs, including basal programs, whole language approaches, and whole word programs.

10. Were studies reporting the largest effects of systematic phonics instruction well designed or poorly designed experiments? That is, was random assignment used? Were the sample sizes sufficiently large? Might results be explained by differences between treatment and control groups that existed prior to the experiment rather than by differences produced by the experimental intervention?

The effects of systematic phonics instruction were not diminished when only the best designed experiments were singled out. The mean effect size for studies using random assignment to place students in treatment and control groups, $\underline{d} = 0.45$, was essentially the same as that for studies employing quasi-experimental designs, $\underline{d} = 0.43$, which utilized existing groups to compare phonics instruction and non-phonics instruction. The mean effect size for studies administering systematic phonics and non-phonics instruction to large samples of students did not differ from studies using the fewest

students: for studies using between 80 and 320 students, $\underline{d}$ = 0.49; for studies using between 20 and 31 students, $\underline{d}$ = 0.48. There were some studies that did not use random assignment and either failed to address the issue of pre-existing differences between treatment and control groups or mentioned that a difference existed but did not adjust for differences in their analysis of results. The effect sizes changed very little when these comparisons were removed from the database, from $\underline{d}$ = 0.44 to $\underline{d}$ = 0.46.

The conclusion drawn is that the significant effects produced by systematic phonics instruction on children's growth in reading were evident in the most rigorously designed experiments. Significant effects did not arise primarily from the weakest studies.

11. Is enough known about systematic phonics instruction to make recommendations for classroom implementation? If so, what cautions should be kept in mind by teachers implementing phonics instruction?

Findings of the panel regarding the effectiveness of systematic phonics instruction were derived from studies conducted in many classrooms with typical classroom teachers and typical American or English-speaking students from a variety of backgrounds and SES levels. Thus, the results of the analysis are indicative of what can be accomplished when systematic phonics programs are implemented in today's classrooms. Systematic phonics instruction has been used widely over a long period with positive results. A variety of phonics programs have proven effective with children of different ages, abilities, and SES backgrounds. These facts should persuade educators and the public that systematic phonics instruction is a valuable part of a successful classroom reading program. The Panel's findings summarized above serve to illuminate the conditions that make systematic phonics instruction especially effective. However, caution is needed in giving a blanket endorsement to all kinds of phonics instruction.

It is important to recognize that the goals of phonics instruction are to provide children with some key knowledge and skills and to ensure that they know how to apply this knowledge in their reading and writing. Phonics teaching is a means to an end. To be able to make use of letter-sound information, children need phonemic awareness. That is, they need to be able to blend sounds together to decode words, and they need to break spoken words into their constituent sounds to write words. Programs that focus too much on the teaching of letter-sounds relations and not enough on putting them to use are unlikely to be very effective. In implementing systematic phonics instruction, educators must keep the *end* in mind and ensure that children understand the purpose of learning letter-sounds and are able to apply their skills in their daily reading and writing activities.

In addition to this general caution, several particular concerns should be taken into consideration to avoid misapplication of the findings. One concern relates to the commonly heard call for "intensive, systematic" phonics instruction. Usually the term "intensive" is not defined, so it is not clear how much teaching is required to be considered intensive. Questions needing further answers are: How many months or years should a phonics program continue? If phonics has been taught systematically in kindergarten and 1st grade, should it continue to be emphasized in 2nd grade and beyond? How long should single instructional sessions last? How much ground should be covered in a program? That is, how many letter-sound relations should be taught and how many different ways of using these relations to read and write words should be practiced for the benefits of phonics to be maximum? These are among the many questions that remain for future research.

Second, the role of the teacher needs to be better understood. Some of the phonics programs showing large effect sizes are scripted so that teacher judgment is largely eliminated. Although scripts may standardize instruction, they may reduce teachers' interest in the teaching process or their motivation to teach phonics. Thus, one concern is how to maintain consistency of instruction and at the same time encourage unique contributions from teachers. Another concern involves what teachers need to know. Some systematic phonics programs require a sophisticated understanding of spelling, structural linguistics, and word etymology. Teachers who are handed the programs but are not provided with sufficient inservice training to use these programs effectively may become frustrated. In view of the evidence showing the effectiveness of systematic phonics instruction, it is important to ensure that the issue of how best to prepare teachers to carry out this

teaching effectively and creatively is given high priority. Knowing that all phonics programs are not the same brings with it the implication that teachers must themselves be educated about how to evaluate different programs and to determine which are based on strong evidence and how they can most effectively use these programs in their own classrooms.

As with any instructional program, there is always the question: "Does one size fit all?" Teachers may be expected to use a particular phonics program with their class, yet it quickly becomes apparent that the program suits some students more than others. In the early grades, children are known to vary greatly in the skills they bring to school. There will be some children who already know most letter-sound correspondences, some children who can even decode words, and others who have little or no letter knowledge. Should teachers proceed through the program and ignore these students? Or should they assess their students' needs and select the types and amounts of phonics suited to those needs? Although the latter is clearly preferable, this requires phonics programs that provide guidance in how to place students into flexible instructional groups and how to pace instruction. However, it is common for many phonics programs to present a fixed sequence of lessons scheduled from the beginning to the end of the school year.

Finally, it is important to emphasize that systematic phonics instruction should be integrated with other reading instruction to create a balanced reading program. Phonics instruction is never a total reading program. In 1st grade, teachers can provide controlled vocabulary texts that allow students to practice decoding, and they can also read quality literature to students to build a sense of story and to develop vocabulary and comprehension. Phonics should not become the dominant component in a reading program, neither in the amount of time devoted to it nor in the significance attached. It is important to evaluate children's reading competence in many ways, not only by their phonics skills but also by their interest in books and their ability to understand information that is read to them. By emphasizing all of the processes that contribute to growth in reading, teachers will have the best chance of making every child a reader.

Directions for Further Research

Although phonics instruction has been the subject of a great deal of study, there are certain extremely important topics that have received little or no research attention, and there are other topics that, although previously studied, require further research to refine our understanding.

Neglected Topics

Three important but neglected questions are prime candidates for research:

(1) What are the "active ingredients" in effective systematic phonics programs? (2) Is phonics instruction improved when motivational factors are taken into account—not only learners' motivation to learn but also teachers' motivation to teach? (3) How does the use of decodable text as early reading material contribute to the effectiveness of phonics programs?

1. Active Ingredients

Systematic phonics programs—even those of the same type, such as synthetic phonics programs—vary in many respects, as indicated in the Panel's report above. It is important to determine whether some properties are essential and others are not. Because instructional time during the school day is limited, teachers and publishers of beginning reading programs need to know which ingredients of phonics programs yield the most benefit. One example of this line of questions involves the content covered. It is clear that the major letter-sound correspondences, including short and long vowels and digraphs, need to be taught. However, there are other regularities of English as well. How far should instruction extend in teaching all of these potential regularities explicitly? Should children be taught to state regularities, or should emphasis be placed on application in reading and writing activities? To what extent do mnemonic devices such as those used in Jolly Phonics (Lloyd, 1993) and Letterland (Wendon, 1992) speed up the process of learning letter shapes, sounds, and names and facilitate their application in reading and writing? What contribution is made by the inclusion of special markings added to written words to clarify how they should be decoded? Research investigating not only these ingredients of phonics programs but other

ingredients as well is needed. These studies should include systematic observation in classrooms to record and analyze the activities of teachers and children using the programs.

2. Motivation

Phonics instruction has often been portrayed as involving "dull drill" and "meaningless worksheets." Such characterizations may accurately describe aspects of some phonics programs, even "effective" ones. Few if any studies have investigated the contribution of motivation to the effectiveness of phonics programs, not only the learner's motivation to learn but also the teacher's motivation to teach. It seems self-evident that the specific techniques and activities used to develop children's letter-sound knowledge and its use in reading and writing should be as relevant and motivating as possible to engage children's interest and attention to promote optimal learning. Moreover, it seems obvious that when the teaching techniques presented to teachers in a phonics program are not only effective but also engaging and enjoyable, teachers will be more successful in their ability to deliver phonics instruction effectively. The lack of attention to motivational factors by researchers in the design of phonics programs is potentially very serious because debates about reading instruction often boil down to concerns about the "relevance" and "interest value" of how something is being taught, rather than the specific content of what is being taught. Future research on phonics instruction should investigate how best to motivate children in classrooms to learn the letter-sound associations and to apply that knowledge to reading and writing. It should also be designed to determine which approaches teachers prefer to use and are most likely to use effectively in their classroom instruction.

3. Decodable Text

Some systematic phonics programs are designed so that children are taught letter-sound correspondences and then provided with little books written carefully to contain the letter-sound relations that were taught. Some programs begin with a very limited set and expand these gradually. The intent of providing books that match children's letter-sound knowledge is to enable them to experience success in decoding words that follow the patterns they know. The stories in such books often involve pigs doing jigs and cats in hats. Other systematic phonics programs make little or no

use of decodable books and select the beginning reading material on some other basis. Some educators reject decodable books outright as too stilted and boring. Surprisingly, very little research has attempted to determine whether the use of decodable books in systematic phonics programs has any influence on the progress that some or all children make in learning to read.

Other Important Topics

The findings of the Panel indicated that systematic phonics instruction provides beginning readers, at-risk readers, disabled readers, and low-achieving readers with a substantial edge in learning to read over alternative forms of instruction not focusing at all or only incidentally on the alphabetic system. However, studies in the database were insufficient in number or in design to address several important satellite questions about the effects of phonics instruction.

Some programs teach many letter-sound relations before children begin using them while other programs introduce a few and then provide reading and writing activities that allow children to apply the correspondences they have learned right away. The latter approach would appear to be preferable, but is it? In what ways does earlier application facilitate growth in reading and writing?

Programs differ in how much time is consumed teaching alphabetic knowledge and word-reading skills. It is unclear how long phonics instruction should continue through the grades. A few studies in the Panel's database indicated that large effect sizes were produced and maintained in the 2nd and 3rd years of instruction for children who were at risk for future reading problems and who began receiving systematic phonics instruction in kindergarten or 1st grade (Blachman et al., (1999); Brown & Felton, 1990; Torgesen et al., 1999). See Table 4 (Appendix E). This suggests that systematic phonics instruction should extend from kindergarten to 2nd grade, but the question remains whether additional instruction will produce further benefits.

It will also be critical to objectively determine the ways in which systematic phonics instruction can be optimally incorporated and integrated in complete and balanced programs of reading instruction. Part of this effort

should be directed at preservice and inservice education to provide teachers with decisionmaking frameworks to guide their selection, integration, and implementation of phonics instruction within a complete reading program.

Another line of questions for research centers around older children above 1st grade who have acquired some reading ability but are reading substantially below grade level. When systematic phonics instruction is introduced to these children, do they have difficulty acquiring alphabetic knowledge and decoding strategies because they have already learned other ways to process print that undermine the acquisition and incorporation of these new processes into their reading? If so, perhaps special steps are required to address this problem. A related question is how can systematic phonics instruction be made more effective for low-achieving readers who have below-average intelligence as well as reading problems. Perhaps instruction in decoding needs to be combined with instruction in reading comprehension strategies to remediate their reading problems.

When systematic phonics instruction is introduced to children who have already acquired some reading skill as a result of another program that does not emphasize phonics, one wonders about the impact of attempting to teach students new strategies when old tricks have already been learned. Findings of the Panel indicated that the impact of systematic phonics instruction was much reduced among children who were introduced to it presumably for the first time in 2nd grade and above. (This presumption may not be accurate, however, because most studies did not state what kind of instruction children had already experienced.) Additional research is needed to study how systematic phonics instruction is received by children who are already reading; whether there are sources of conflict; and, if so how to address them instructionally. A related question is whether the sequence of instruction makes a difference. It may be that children do better when a year of systematic phonics instruction precedes a year of whole language instruction than when the reverse is the case.

References

Adams, M. J. (1990). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Anderson, R. C., Hiebert, E. F., Wilkinson, I. A. G., & Scott, J. (1985). Becoming a nation of readers. Champaign, IL: Center for the Study of Reading.

Aukerman, R. C. (1971). Approaches to beginning reading. New York: John Wiley & Sons.

Aukerman, R. C. (1981). The basal reader approach to reading. New York: John Wiley & Sons.

Aukerman, R. C. (1984). Approaches to beginning reading (2nd ed.). New York: John Wiley & Sons.

Balmuth, M. (1982). The roots of phonics: A historical introduction. New York: McGraw-Hill.

Barr, R., Kamil, M., Mosenthal, P., & Pearson, P. (1991). Handbook of reading research (Vol. 2). New York: Longman.

Baumann, J., Hoffman, J., Moon, J., & Duffy-Hester, A. (1998). Where are teachers' voices in the phonics/whole language debate? Results from a survey of U.S. elementary classroom teachers. The Reading Teacher, 51, 636-650.

Bear, D. R., Invernizz, M., Templeton, S., & Johnson, F. (2000). Words their way: Word study for phonics, vocabulary and spelling instruction. Upper Saddle River, NJ: Merrill.

Beck, I., & Mitroff, D. (1972). The rationale and design of a primary grades reading system for an individualized classroom. Pittsburgh: University of Pittsburgh Learning Research and Development Center.

Blachman, B. (in press). Phonological awareness. In M. Kamil, P. Mosenthal, P. Pearson, & R. Barr (Eds.), Handbook of reading research (Vol. 3). Mahwah, NJ: Erlbaum.

Bond, G., & Dykstra, R. (1967/1998). The cooperative research program in first grade reading. Reading Research Quarterly, 2, 5-142.

Brown, I., & Felton, R. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. Reading and Writing: An Interdisciplinary Journal, 2, 223-241.

Bruck, M. (1993). Component spelling skills of college students with childhood diagnoses of dyslexia. Learning Disability Quarterly, 16, 171-184.

Campbell, D., & Stanley, J. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

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

Chall, J. S. (1967). Learning to read: The great debate (1st ed.). New York, NY: McGraw-Hill.

Chall, J. S. (1983). Learning to read: The great debate (2nd ed.). New York, NY: McGraw-Hill.

Chall, J. (1992). The new reading debates: Evidence from science, art, and ideology. Teachers College Record, 94, 315-328.

Chall, J. (1996a). Learning to read: The great debate (revised, with a new foreword). New York: McGraw-Hill.

Chall, J. (1996b). Stages of reading development (2nd ed.). Fort Worth, TX: Harcourt-Brace.

Chall, J., & Feldmann., S. (1966). First grade reading: An analysis of the interactions of professed methods, teacher implementation, and child background. The Reading Teacher, 19, 569-575.

Clay, M. M. (1993). Reading recovery: A guidebook for teachers in training. Portsmouth, NH: Heinemann.

Cohen, J. (1988). Statistical power analysis for the behavior sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Cox, A. R. (1991). Structures and techniques: Multisensory teaching of basic language skills. Cambridge, MA: Educators Publishing Service.

Dahl, K. L., Sharer, P. L., Lawson, L. L., & Grogran, P. R. (1999). Phonics instruction and student achievement in whole language first-grade classrooms. Reading Research Quarterly, 34, 312-341.

Dickson, S. (1972a). Sing, spell, read, & write. Chesapeake, VA: International Learning Systems.

Dykstra, R. (1968). The effectiveness of code- and meaning-emphasis in beginning reading programs. The Reading Teacher, 22, 17-23.

Ehri, L. (1991). Development of the ability to read words. In R. Barr, M. Kamil, P. Mosenthal, & P. Pearson (Eds.), Handbook of reading research (Vol. 2) (pp. 383-417). New York: Longman.

Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. Ehri, & R. Treiman (Eds.), Reading acquisition (pp. 107-143). Hillsdale, NJ: Erlbaum.

Ehri, L. (1994). Development of the ability to read words: Update. In R. Ruddell, M. Ruddell, & H. Singer (Eds.), Theoretical models and processes of reading (4th ed. pp. 323-358). Newark, DE: International Reading Association.

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), Word recognition in beginning literacy (pp. 3-40). Mahwah, NJ: Erlbaum.

Ehri, L., Deffner, N., & Wilce, L. (1984). Pictorial mnemonics for phonics. Journal of Educational Psychology, 76, 880-893.

Engelmann, S. (1980). Direct instruction. Englewood Cliffs, NJ: Prentice-Hall.

Engelmann, S., & Bruner, E. (1969). Distar reading program. Chicago: SRA.

Engelmann, S., & Bruner, E. (1978). DISTAR reading: An instructional system. Chicago: Science Research Associates.

Engelmann, S., & Bruner, E. (1988). Reading mastery I/II fast cycle: Teacher's guide. Chicago: Science Research Associates, Inc.

Engelmann, S., & Osborn, J. (1987). DISTAR Language I. Chicago: Science Research Associates.

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), Word recognition in beginning literacy (pp. 3-40). Mahwah, NJ: Erlbaum.

Ehri, L., Deffner, N., & Wilce, L. (1984). Pictorial mnemonics for phonics. Journal of Educational Psychology, 76, 880-893.

Engelmann, S. (1980). Direct instruction. Englewood Cliffs, NJ: Prentice-Hall.

Engelmann, S., & Bruner, E. (1969). Distar reading program. Chicago: SRA.

Engelmann, S., & Bruner, E. (1978). DISTAR reading: An instructional system. Chicago: Science Research Associates.

Engelmann, S., & Bruner, E. (1988). Reading mastery I/II fast cycle: Teacher's guide. Chicago: Science Research Associates, Inc.

Engelmann, S., & Osborn, J. (1987). DISTAR Language I. Chicago: Science Research Associates.

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), Word recognition in beginning literacy (pp. 3-40). Mahwah, NJ: Erlbaum.

Ehri, L., Deffner, N., & Wilce, L. (1984). Pictorial mnemonics for phonics. Journal of Educational Psychology, 76, 880-893.

Engelmann, S. (1980). Direct instruction. Englewood Cliffs, NJ: Prentice-Hall.

Engelmann, S., & Bruner, E. (1969). Distar reading program. Chicago: SRA.

Engelmann, S., & Bruner, E. (1978). DISTAR reading: An instructional system. Chicago: Science Research Associates.

Engelmann, S., & Bruner, E. (1988). Reading mastery I/II fast cycle: Teacher's guide. Chicago: Science Research Associates, Inc.

Engelmann, S., & Osborn, J. (1987). DISTAR Language I. Chicago: Science Research Associates.

Foorman, B., Francis, D., Fletcher, J., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. Journal of Educational Psychology, 90, 37-55.

Freppon, P. A., & Dahl, K. L. (1991). Learning about phonics in a whole language classroom. Language Arts, 68, 190-197.

Freppon, P. A., & Headings, L. (1996). Keeping it whole in whole language: A first grade teacher's instruction in an urban whole language classroom. In E. McIntyre & M. Pressley (Eds.), Balanced instruction: Strategies and skills in whole language (pp. 65-82). Norwood, MA: Christopher-Gordon.

Gaskins, I., Downer, M., & Gaskins, R. (1986). Introduction to the Benchmark school word identification/vocabulary development program. Media, PA: Benchmark School.

Gilllingham, A., & Stillman, B. (1940). Remedial training for children with specific disability in reading, spelling, and penmanship. New York, NY: Sackett & Williams Lithographing Co.

Gillingham, A., & Stillman, B. W. (1979). Remedial training for children with specific disability in reading, spelling, and penmanship. Cambridge, MA: Educators Publishing Service.

Goodman, K. S. (1993). Phonics phacts. Portsmouth, NH: Heinemann.

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

Greaney, K., Tunmer, W., & Chapman, J. (1997). Effects of rime-based orthographic analogy training on the word recognition skills of children with reading disability. Journal of Educational Psychology, 89, 645-651.

Grundin, H. U. (1994). If it ain't whole, it ain't language—or back to the basics of freedom and dignity. In F. Lehr & J. Osborn (Eds.), Reading, language, and literacy (pp. 77-88). Mahwah, NJ: Erlbaum.

Harris, T., & Hodges, R. (Eds.). (1995). The literacy dictionary. Newark, DE: International Reading Association.

Haskell, D., Foorman, B., & Swank, P. (1992). Effects of three orthographic/phonological units on first-grade reading. Remedial and Special Education, 13, 40-49.

Holdaway, D. (1979). The foundations of literacy. Sydney: Ashton-Scholastic.

Hoover, W., & Gough, P. (1990). The simple view of reading. Reading and Writing, 2, 127-160.

Iversen, S., & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery Program. Journal of Educational Psychology, 85, 112-126.

Johnson, B. (1989). DSTAT: Software for the meta-analytic review of research literatures. Hillsdale, NJ: Erlbaum.

Juel, C. (1994). Learning to read and write in one elementary school. New York: Springer-Verlag.

Kameenui, E. J., Simmons, D. C., Chard, D., & Dickson, S. (1997). Direct instruction reading. In S. A. Stahl & D. A. Hayes (Eds.), Instructional models in reading (pp. 59-84). Mahwah, NJ: Erlbaum.

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

Leach, D., & Siddall, S. (1990). Parental involvement in the teaching of reading: A comparison of hearing reading, paired reading, pause, prompt, praise, and direct instruction methods. British Journal of Educational Psychology, 60, 349-355.

Lindamood, C. H., & Lindamood, P. C. (1984). Auditory discrimination in depth. Austin, TX: PRO-ED, Inc.

Lloyd, S. (1993). The phonics handbook. UK: Jolly Learning.

Levin, H., & Williams, J. (Eds.). (1970). Basic studies on reading. New York: Basic Books.

Lovett, M., Borden, S., DeLuca, T., Lacerenza, L., Benson, N., & Brackstone, D. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. Developmental Psychology, 30, 805-822.

Lovett, M., Lacerenza, L., Borden, S., Frijters, J., Steinbach, K., & DePalma, M. (in press). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. Journal of Educational Psychology.

Lovett, M., & Steinbach, K. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children? Learning Disability Quarterly, 20, 189-210.

Lovett, M., Warren-Chaplin, P., Ransby, M., & Borden, S. (1990). Training the word recognition skills of reading disabled children: Treatment and transfer effects. Journal of Educational Psychology, 82, 769-780.

Lovitt, T. C., & DeMier, D. M. (1984). An evaluation of the Slingerland method with LD youngsters. Journal of Learning Disabilities, 17, 267-272.

McCracken, G., & Walcutt, C. (1963, 1965, 1966, 1975). Basic reading. Philadelphia, PA: Lippincott.

McKenna, M., Stahl, S., & Reinking, D. (1994). A critical commentary on research, politics, and whole language. Journal of Reading Behavior, 26, 211-233.
Lovitt, T. C., & DeMier, D. M. (1984). An evaluation of the Slingerland method with LD youngsters. Journal of Learning Disabilities, 17, 267-272.

McCracken, G., & Walcutt, C. (1963, 1965, 1966, 1975). Basic reading. Philadelphia, PA: Lippincott.

McKenna, M., Stahl, S., & Reinking, D. (1994). A critical commentary on research, politics, and whole language. Journal of Reading Behavior, 26, 211-233.

Meyer, L. A. (1983). Increased student achievement in reading: One district's strategies. Research in Rural Education, 1, 47-51.

Mills, H., O'Keefe, T., & Stephens, D. (1992). Looking closely: Exploring the role of phonics in one whole language classroom. Urbana, IL: National Council of Teachers of English.

Morris, D. (1992). Case studies in teaching beginning readers: The Howard Street tutoring manual. Boone, NC: Fieldstream.

Morris, D., & Perney, J. (1984). Developmental spelling as a predictor of first grade reading achievement. Elementary School Journal, 84, 441-457.

Ogden, S., Hindman, S., & Turner, S. (1989). Multisensory programs in the public schools: A brighter future for LD children. Annals of Dyslexia, 39, 247-267.

Pearson, P., Barr, R., Kamil, M., & Mosenthal, P. (1984). Handbook of reading research. New York: Longman.

Popp, H. M. (1975). Current practices in the teaching of beginning reading. In J. B. Carroll & J. S. Chall (Eds.), Toward a literate society (pp. 101-146). New York: McGraw-Hill.

Purves, A. (Ed.). (1994). Encyclopedia of English studies and language arts. New York: Scholastic.

Rack, J., Snowling, M., & Olson, R. (1992). The nonword reading deficit in developmental dyslexia: A review. Reading Research Quarterly, 27, 29-53.

Routman, R. (1996). Literacy at the crossroads. Portsmouth, NH: Heinemann.

Santa, C., & Hoien, T. (1999). An assessment of early steps: A program for early intervention of reading problems. Reading Research Quarterly, 34, 54-79.

Shankweiler, D., & Liberman, I. Y. (1972). Misreading: A search for causes. In J. F. Kavanaugh & I. G. Mattingly (Eds.), Language by eye and by ear (pp. 293-317). Cambridge, MA: MIT Press.

Share, D. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. Cognition, 55, 151-218.

Share, D., Jorm, A., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading achievement. Journal of Educational Psychology, 76, 1309-1324.

Snider, V. (1990). Direct instruction reading with average first-graders. Reading Improvement, 27, 143-148.

Stahl, S. A. (1999). Why innovations come and go: The case of whole language. Educational Researcher.

Stahl, S. A., Duffy-Hester, A. M., & Stahl, K. A. D. (1998). Everything you wanted to know about phonics (but were afraid to ask). Reading Research Quarterly, 35, 338-355.

Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. Review of Educational Research, 59 (1), 87-116.

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

Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics teaching improves rading and spelling in inner-city second language learners. British Journal of Educational Psychology, 69, 587-605.

Tangel, D., & Blachman, B. (1995). Effect of phoneme awareness instruction on the invented spelling of first-grade children: A one-year follow-up. Journal of Reading Behavior, 27, 153-185.

Taylor, D. (1998). Beginning to read & the spin doctors of science: The political campaign to change America's mind about how children learn to read. Urbana, IL: National Council of Teachers of English.

Torgesen, J., Wagner, R., Rashotte, C., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. Journal of Educational Psychology, 91, 579-593.

Tunmer, W., & Chapman, J. (1998). Language prediction skill, phonological recoding ability, and beginning reading. In C. Hulme & R. Joshi (Eds.), Reading and spelling: Development and disorders (pp. 33-67). Mahwah, NJ: Erlbaum.

Tunmer, W., & Hoover, W. (1993). Phonological recoding skill and beginning reading. Reading and Writing: An Interdisciplinary Journal, 5, 161-179.

Vandervelden, M., & Siegel, L. (1997). Teaching phonological processing skills in early literacy: A developmental approach. Learning Disability Quarterly, 20, 63-81.

Vickery, K., Reynolds, V., & Cochran, S. (1987). Multisensory teaching approach for reading, spelling, and handwriting, Orton-Gillingham based curriculum, in a public school setting. Annals of Dyslexia, 37, 189-200.

Weaver, C. (1998). Experimental research: On phonemic awareness and on whole language. In C. Weaver (Ed.), Reconsidering a balanced approach to reading (pp. 321-371). Urbana, IL: National Council of Teachers of English.

Wendon, L. (1992). First steps in Letterland. Cambridge, UK: Letterland Ltd.

Woodcock, R. (1987). Woodcock reading mastery tests - revised. Circle Pines, MN: American Guidance Service.

PART II: PHONICS INSTRUCTION
Appendices

Appendix A
Studies Included in the Meta-Analysis

Note: Studies were assigned numbers during the screening process. Numbers missing in the list were assigned to studies that were rejected for the analysis. These are listed separately.

03 Blachman, B., Tangel, D., Ball, E., Black, R. & McGraw, D. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. Reading and Writing: An Interdisciplinary Journal, 11, 293-273.

04 Bond, C., Ross, S., Smith, L., & Nunnery, J. (1995-96). The effects of the sing, spell, read, and write program on reading achievement of beginning readers. Reading Research and Instruction, 35, 122-141.

05 Brown, I., & Felton, R. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. Reading and Writing: An Interdisciplinary Journal, 2, 223-241.

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

09 Evans, M., & Carr, T. (1985). Cognitive abilities, conditions of learning, and the early development of reading skill. Reading Research Quarterly, 20, 327-350.

11 Foorman, B., Francis, D., Fletcher, J., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. Journal of Educational Psychology, 90, 37-55.

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

13 Foorman, B., Francis, D., Winikates, D., Mehta, P., Schatschneider, C., & Fletcher, J. (1997). Early interventions for children with reading disabilities. Scientific Studies of Reading, 1, 255-276.

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

15 Fulwiler, G., & Groff, P. (1980). The effectiveness of intensive phonics. Reading Horizons, 21, 50-54.

72 Gersten, R., Darch, C., & Gleason, M. (1988). Effectiveness of a direct instruction academic kindergarten for low-income students. The Elementary School Journal, 89, 227-240.

17 Gittelman, R., & Feingold, I. (1983). Children with reading disorders—I. Efficacy of reading remediation. Journal of Child Psychology and Psychiatry and Allied Disciplines, 24, 167-191.

18 Greaney, K., Tunmer, W., & Chapman, J. (1997). Effects of rime-based orthographic analogy training on the word recognition skills of children with reading disability. Journal of Educational Psychology, 89, 645-651.

60 Griffith, P., Klesius, J., & Kromey, J. (1992). The effect of phonemic awareness on the literacy development of first grade children in a traditional or a whole language classroom. Journal of Research in Childhood Education, 6, 85-92.

22 Haskell, D., Foorman, B., & Swank, P. (1992). Effects of three orthographic/phonological units on first-grade reading. Remedial and Special Education, 13, 40-49.

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

28 Leach, D., & Siddall, S. (1990). Parental involvement in the teaching of reading: A comparison of hearing reading, paired reading, pause, prompt, praise, and direct instruction methods. British Journal of Educational Psychology, 60, 349-355.

29 Leinhardt, G., & Engel, M. (1981). An iterative evaluation of NRS: Ripples in a pond. Evaluation Review, 5, 579-601.

75 Lovett, M., Lacerenza, L., Borden, S., Frijters, J., Steinbach, K., & DePalma, M. (in press). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. Journal of Educational Psychology.

32 Lovett, R., Ransby, M., Hardwick, N., Johns, M., & Donaldson, S. (1989). Can dyslexia be treated? Treatment-specific and generalized treatment effects in dyslexic children's reponse to remediation. Brain and Language, 37, 90-121.

33 Lovett, M., & Steinbach, K. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children? Learning Disability Quarterly, 20, 189-210.

34 Lovett, M., Warren-Chaplin, P., Ransby, M., & Borden, S. (1990). Training the word recognition skills of reading disabled children: Treatment and transfer effects. Journal of Educational Psychology, 82, 769-780.

35 Lum, T., & Morton, L. (1984). Direct instruction in spelling increases gain in spelling and reading skills. Special Education in Canada, 58, 41-45.

36 Mantzicopoulos, P., Morrison, D., Stone, E., & Setrakian, W. (1992). Use of the SEARCH/TEACH tutoring approach with middle-class students at risk for reading failure. Elementary School Journal, 92, 573-586.

37 Marston, D., Deno, S., Kim, D., Diment, K., & Rogers, D. (1995). Comparison of reading intervention approaches for students with mild disabilities. Exceptional Children, 62, 20-37.

38 Martinussen, R., & Kirby, J. (1998). Instruction in successive and phonological processing to improve the reading acquisition of at-risk kindergarten children. Developmental Disabilities Bulletin, 26, 19-39.

41 Oakland, T., Black, J., Stanford, G., Nussbaum, N., & Balise, R. (1998). An evaluation of the dyslexia training program: A multisensory method for promoting reading in students with reading disabilities. Journal of Learning Disabilities, 31, 140-147.

44 Santa, C., & Hoien, T. (1999). An assessment of early steps: A program for early intervention of reading problems. Reading Research Quarterly, 34, 54-79.

47 Silberberg, N., Iversen, I., & Goins, J. (1973). Which remedial reading method works best? Journal of Learning Disabilities, 6, 18-27.

48 Snider, V. (1990). Direct instruction reading with average first-graders. Reading Improvement, 27, 143-148.

74 Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics teaching improves rading and spelling in inner-city second language learners. British Journal of Educational Psychology, 69, 587-605.

51   Torgesen, J., Wagner, R., Rashotte, C., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. Journal of Educational Psychology, 91, 579-593.

52   Traweek, K., & Berninger, V. (1997). Comparisons of beginning literacy programs: Alternative paths to the same learning outcome. Learning Disability Quarterly, 20, 160-168.

53   Tunmer, W., & Hoover, W. (1993). Phonological recoding skill and beginning reading. Reading and Writing: An Interdisciplinary Journal, 5, 161-179.

69   Umbach, B., Darch, C., & Halpin, G. (1989). Teaching reading to low performing first graders in rural schools: A comparison of two instructional approaches. Journal of Instructional Psychology, 16, 23-30.

54   Vandervelden, M., & Siegel, L. (1997). Teaching phonological processing skills in early literacy: A developmental approach. Learning Disability Quarterly, 20, 63-81.

55   Vickery, K., Reynolds, V., & Cochran, S. (1987). Multisensory teaching approach for reading, spelling, and handwriting, Orton-Gillingham based curriculum, in a public school setting. Annals of Dyslexia, 37, 189-200.

57   Wilson, K., & Norman, C. (1998). Differences in word recognition based on approach to reading instruction. Alberta Journal of Educational Research, 44, 221-230.

Appendix B

List of Studies in the Original Database That Were Excluded

62   Barr, R. (1972). The influence of instructional conditions on word recognition errors. Reading Research Quarterly, 7 (3), 509-529.

63   Barr, R. (1974). The effect of instruction on pupil reading strategies. Reading Research Quarterly, 10, (4), 555-582.

01   Becker, W., & Gersten, R. (1982). A follow-up of follow through: The later effects of the direct instruction model on children in fifth and sixth grades. American Educational Research Journal, 19(1), 75-92.

02   Berninger, V., Vaughan, K., Abbott, R., Brooks, A., Abbott, S., Rogan, L., Reed, E., & Graham, S. (1998). Early intervention for spelling problems: Teaching functional spelling units of varying size with a multiple-connections framework. Journal of Educational Psychology, 90, 587-605.

61   Blachowicz, C., McCarthy, L., & Ogle, D. (1979). Testing phonics: A look at children's response biases. Illinois School Research and Development, 16 (1), 1-6.

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

71   Carnine, D. (1980). Phonic versus whole-word correction procedures following phonic instruction. Education and Treatment of Children, 3(4), 323-329.

06   Eddowes, E. (1990). Teaching reading in kindergarten: Two contrasting approaches. Reading Improvement, 27, 220-223.

07   Ehri, L., & Wilce, L. (1987). Cipher versus cue reading: An experiment in decoding acquisition. Journal of Educational Psychology, 79 (1), 3-13.

10   Fielding-Barnsley, R (1997). Explicit instruction in decoding benefits children high in phonemic awareness and alphabet knowledge. Scientific Studies of Reading, 1, 85-98.

14   Foorman, B., & Liverman, D. (1989). Visual and phonological processing of words: A comparison of good and poor readers. Journal of Learning Disabilities, 22, 349-355.

66   Gersten, R., Keating, T., & Becker, W. (1988). The continued impact of the direct instruction model: Longitudinal studies of follow through students. Education and Treatment of Children, 11 (4), 318-327.

16   Gettinger, M. (1987). Prereading skills and achievement under three approaches to teaching word recognition. Journal of Research and Development in Education, 19, 1-9.

68   Gillon, G., & Dodd, B. (1997). Enhancing the phonological processing skills of children with specific reading disability. European Journal of Disorders of Communication, 32, 67-90.

19   Guyer, B., Banks, S., & Guyer, K. (1993). Spelling improvement for college students who are dyslexic. Annals of Dyslexia, 43, 186-193.

20   Guyer, B., & Sabatino, D. (1989). The effectiveness of a multisensory alphabetic phonetic approach with college students who are learning disabled. Journal of Learning Disabilities, 22, 430-434.

21   Hall, D., & Cunningham, P. (1988). Context as a polysyllabic decoding strategy. Reading Improvement, 25, 261-264.

23   Jeffrey, W., & Samuels, S. (1967). Effect of method of reading training on initial learning and transfer. Journal of Verbal Learning and Verbal Behavior, 6, 354-358.

73   Kameenui, E., Stein, M., Carnine, D., & Maggs, A. (1981). Primary level word attack skills based on isolated word, discrimination list and rule application training. Reading Education, 6 (2), 46-55.

24 Kaiser, S., Palumbo, K., Bialozor, R., & McLaughlin, T. (1989). The effects of direct instruction with rural remedial students: A brief report. Reading Improvement, 26, 88-93.

25 Kibby, M. (1979). The effects of certain instructional conditions and response modes on initial word learning. Reading Research Quarterly, 15, 147-171.

27 Kuder, S. (1990). Effectiveness of the DISTAR Reading Program for children with learning disabilities. Journal of Learning Disabilities, 23, 69-71.

30 Lovett, M., Borden, S., DeLuca, T., Lacerenza, L., Benson, N., & Brackstone, D. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. Developmental Psychology, 30, 805-822.

31 Lovett, M., Ransby, M., & Barron, R. (1988). Treatment, subtype, and word type effects in dyslexic children's response to remediation. Brain and Language, 34, 328-349.

39 McNeil, J., & Donant, L. (1980). Transfer effect of word recognition strategies. Journal of Reading Behavior, 12(2), 97-103.

40 Meyer, L. (1982). The relative effects of word-analysis and word-supply correction procedures with poor readers during word-attack training. Reading Research Quarterly, 17(4), 544-555.

42 Olson, R., & Wise, B. (1992). Reading on the computer with orthographic and speech feedback. Reading and Writing: An Interdisciplinary Journal, 4, 107-144.

43 Olson, R., Wise, B., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the post-training development of word recognition. Scientific Studies of Reading, 1(3), 235-253.

65 Peterson, M., & Haines, L. (1992). Orthographic analogy training with kindergarten children: Effects on analogy use, phonemic segmentation, and letter-sound knowledge. Journal of Reading Behavior, 24 (1), 109-127.

45 Scherzer, C. & Goldstein, D. (1982). Children's first reading lesson: Variables influencing within-lesson emotional behavior and post-lesson achievement. Journal of Educational Psychology, 74(3), 382-392.

46 Schwartz, S. (1988). A comparison of componential and traditional approaches to training reading skills. Applied Cognitive Psychology, 2, 189-201.

49 Spiegel, D., Fitzgerald, J., & Heck, M. (1985). Teaching children to use a context-plus-phonics strategy. Reading Horizons, 25, 176-185.

50 Sullivan, H., Okada, M., & Niedermeyer, F. (1971). Learning and transfer under two methods of word-attack instruction. American Educational Research Journal, 8(2), 227-239.

67 Torgesen, J., Wagner, R., & Rashotte, C. (1997). Prevention and remediation of severe reading disabilities: Keeping the end in mind. Scientific Studies of Reading, 1 (3), 217-234.

56 Walton, P. (1995). Rhyming ability, phoneme identify, letter-sound knowledge, and the use of orthographic analogy by pre-readers. Journal of Educational Psychology, 87(4), 587-597.

70 Wise, B., Olson, R., Anstett, M., Andrews, L., Terjak, M., Schneider, V., Kostuch, J., & Driho, L. (1989). Implementing a long-term computerized remedial reading program with synthetic speech feedback: Hardware, software, and real-world issues. Behavior Research Methods, Instruments and Computers, 21 (2), 173-180.

58 Wise, B., & Olson, R. (1992). How poor readers and spellers use interactive speech in a computerized spelling program. Reading and Writing: An Interdisciplinary Journal, 4, 145-163.

Appendix C

In the reports of experiments included in the meta-analysis examining the effects of phonics instruction, references were supplied for the programs and material used to teach systematic phonics. These are listed below.

Synthetic Phonics Programs

03   Blachman et al., 1999

Blachman, B. (1987). An alternative classroom reading program for learning disabled and other low achieving children. In R. Bowler (Ed.), Intimacy with language: A forgotten basic in teacher education (pp. 49-55). Baltimore, MD: Orton Dyslexia Society.

Engelmann, S. (1969). Preventing failure in the primary grades. Chicago: Science Research Associates.

Slingerland, B. (1971). A multi-sensory approach to language arts for specific language disability children: A guide for primary teachers. Cambridge, MA: Educators Publishing Service.

Primary phonics series – published by Educator's Publishing Service.

Selected stories from Scott Foresman basal reading series (none of its other materials was used).

04   Bond et al., 1995

Dickson, S. (1972a). Sing, spell, read, & write. Chesapeake, VA: International Learning Systems.

05   Brown & Felton, 1990

Lippincott basic reading. (1981). Riverside, NJ: Macmillan Publishing Company.

11   Foorman et al., 1998

Open Court Reading. (1995). Collections for young scholars. Chicago and Peru, IL: SRA/McGraw-Hill.

12   Foorman et al., 1991

Scott, Foresman (1985). Scott, Foresman reading. Glenview, IL: Author.

Modern Curriculum Press. (1986). Phonics practice readers, series B. Cleveland, OH: Author.

Elwell, C., Murray, P., & Kucia, M. (1988). Modern Curriculum Press phonics program. Cleveland, OH: Modern Curriculum Press.

13   Foorman et al., 1997

Cox, A. R. (1991). Structures and techniques: Multisensory teaching of basic language skills. Cambridge, MA: Educators Publishing Service.

Gillingham, A., & Stillman, B. (1979). Remedial training for children with specific disability in reading, spelling, and penmanship. Cambridge, MA: Educators Publishing Service.

Fulwiler & Groff, 1980

McCracken, G., & Walcutt, C. (1975). Basic reading. Philadelphia: Lippincott.

17   Gittelman & Feingold, 1983

Pollack, C. (1967a). Phonic readiness kit. Brooklyn, NY: Book-Lab, Inc.

Pollack, C. (1967b). The intersensory reading program: My own reading book and my own writing book, part I. Brooklyn, NY: Book-Lab, Inc.

Pollack, C., & Atkins, R. (1971). The intersensory reading program: My own reading book, part II. Brooklyn, NY: Book-Lab, Inc.

Pollack, C., & Lane, P. (1969). The lipreader, part I. Brooklyn, NY: Book-Lab, Inc.

Pollack, C., & Lane, P. (1971) The lipreader, part II. Brooklyn, NY: Book-Lab, Inc.

28   Leach & Sidall, 1990

Engelmann, S. (1980). <u>Direct instruction</u>. Englewood Cliffs, NJ: Prentice-Hall.

Engelmann, S., & Bruner, E. (1975). <u>DISTAR reading I & II: An instructional system</u>. Chicago: Science Research Associates.

Engelmann, S., Haddox, P., & Bruner, E. (1983). <u>Teach your child to read in 100 easy lessons</u>. New York: Simon & Schuster.

29   Leinhardt & Engel, 1981

Beck, I. (1977). Comprehension during the acquisition of decoding skills. In J. T. Guthrie (Ed.), <u>Cognition, curriculum, and comprehension</u>. Newark, DE: International Reading Association.

Beck, I., & Mitroff, D. (1972). <u>The rationale and design of a primary grades reading system for an individualized classroom</u>. Pittsburgh: University of Pittsburgh Learning Research and Development Center.

Lovett & Steinbach, 1997

Engelmann, S., & Bruner, E. C. (1988). <u>Reading mastery I/II fast cycle: Teacher's guide</u>. Chicago: Science Research Associates, Inc.

Engelmann, S., Carnine, L., & Johnson, G. (1988). <u>Corrective reading, word attack basics, decoding A</u>. Chicago: Science Research Associates, Inc.

Engelmann, S., Johnson, G., Carnine, L., Meyer, L., Becker, W., & Eisele, J. (1988). <u>Corrective reading: Decoding strategies, decoding B1</u>. Chicago: Science Research Associates, Inc.

37   Marston et al., 1995

Engelmann, S. (1980). <u>Direct instruction</u>. Englewood Cliffs, NJ: Educational Technology Publications.

Science Research Associates (1988). <u>Reading mastery</u>. New York: SRA, Macmillan/McGraw-Hill.

38   Martinussen & Kirby, 1998

Das, J., Mishra, R., & Pool, J.(1995). An experiment on cognitive remediation of reading difficulties. <u>Journal of Learning Disabilities, 28</u>, 66-79.

41   Oakland et al., 1998

Beckham, P. B., & Biddle, M. L. (1989). <u>Dyslexia training program books</u>. Cambridge, MA: Educators Publishing Service.

48   Snider, 1990

Carnine, D., & Silbert, J. (1979). <u>Direct instruction reading</u>. Columbus, OH: Merrill.

51   Torgesen et al., 1999

Lindamood, C. H., & Lindamood, P. C. (1984). <u>Auditory discrimination in depth</u>. Austin, TX: PRO-ED, Inc.

Short stories from:

Hannah, S. (1993). <u>Early literacy series, I & II</u>. Edmonton, Alberta, Canada: The Centre for Literacy.

Smith, L. (1992) <u>Auditory discrimination reading series-sequence B</u>. Independence, MO: Poppin Creations.

52   Traweek & Berninger, 1997

Engelmann, S., & Bruner, E. C. (1978). <u>DISTAR reading: An instructional system</u>. Chicago: Science Research Associates.

55   Vickery et al., 1987

Gillingham, A., & Stillman, B. W. (1956). <u>Remedial training for children with specific disability in reading, spelling, and penmanship</u> (5th ed). Cambridge, MA: Educators Publishing Service.

Cox, A. (1984). <u>Structures and techniques: Multisensory teaching of basic language skills</u>. Cambridge, MA: Educators Publishing Service.

Waites, L., & Cox, A. (1969). Developmental language disability . . . Basic training . . . Remedial language training. Cambridge, MA: Educators Publishing Service.

69   Umbach et al., 1992

Reading mastery series (1986)

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

72   Gersten et al., 1988

Becker, W., Engelmann, S., Carnine, D., & Rhine, R. (1981). Direct instruction models. In R. Rhine (Ed.), Encouraging change in American schools: A decade of experimentation (pp. 95-154). New York: Academic Press.

Carnine, D., Carnine, L., Karp, J., & Weisbert, P. (1988) Kindergarten for economically disadvantaged students: The direct instruction component. In C. Wargen (Ed.), A resource guide to public school early childhood programs (pp. 73-98). Alexandria, VA: ASCD

Engelmann, S., & Osborn, J. (1987). DISTAR language I. Chicago: Science Research Associates.

74   Stuart, 1999

Lloyd, S. (1992). The phonics handbook. UK: Jolly Learning.

Lovett et al., in press

Engelmann, S., & Bruner, E. C. (1988). Reading mastery I/II fast cycle: Teacher's guide. Chicago: Science Research Associates, Inc.

Engelmann, S., Carnine, L., & Johnson, G. (1988). Corrective reading, word attack basics, decoding A. Chicago: Science Research Associates, Inc.

Engelmann, S., Johnson, G., Carnine, L., Meyer, L., Becker, W., & Eisele, J. (1988). Corrective reading: Decoding strategies, decoding B1. Chicago: Science Research Associates, Inc.

Phonics Programs Emphasizing Larger Phonological Subunits

11   Foorman et al., 1998

Hiebert, E., Colt, J., Catto, S., & Gary, E. (1992). Reading and writing of first-grade students in a restructured chapter 1 program. American Educational Research Journal, 29, 545-572.

13   Foorman et al., 1997

Traub, N., & Bloom, F. (1992). Recipe for reading (3rd ed.). Cambridge, MA: Educators Publishing Service.

33   Lovett & Steinbach, 1997

Gaskins, I., Downer, M., & Gaskins, R. (1986). Introduction to the Benchmark school word identification/vocabulary development program. Media, PA: Benchmark School.

Gaskins, I., Downer, M., Anderson, R., et al. (1988). A metacognitive approach to phonics: Using what you know to decode what you don't know. Remedial and Special Education, 9, 36-41, 66.

44   Santa & Hoien, 1999

Morris, D., Shaw, B., & Perney, J. (1990). Helping low readers in grades 2 and 3: An after school volunteer program. Elementary School Journal, 91, 133-150.

Santa, C. (1998). Early steps: Learning from a reader. Kalispell, MT: Scott.

51   Torgesen et al., 1999

Short stories from HBJ bookmark Series

53   Tunmer & Hoover, 1993

Clay , M. (1985). The early detection of reading difficulties. Auckland: Heinemann.

Clay, M. (1991). Becoming literate: The construction of inner control. Auckland: Heinemann.

Iversen, S., & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery program. Journal of Educational Psychology, 85, 112-126.

75   Lovett et al., in press

Gaskins, I., Downer, M., & Gaskins, R. (1986). Introduction to the Benchmark school word identification/vocabulary development program. Media, PA: Benchmark School.

Miscellaneous Phonics Programs:

34   Lovett et al., 1990

Dixon, R., Engelmann, S., & Meier, M. (1980). Spelling mastery, a direction instruction series - Level A. Chicago: Science Research Asssociates.

Dixon, R., & Engelmann, S. (1981). Spelling mastery, a direction instruction series - Level B. Chicago: Science Research Associates.

Mantzicopoulos et al., 1992

Cradler, J., Bechthold, D., & Bechthold, C. (1974). Success-controlled optimal reading experience (SCORE). Hillsborough, CA: Learning Guidance Systems.

Makar, B. (1985). Primary phonics. Cambridge, MA: Educators Publishing Service.

Phonics practice readers. (1986). Cleveland, OH: Modern Curriculum Press.

Hall & Price (1984). Explode the code. Cambridge, MA: Educators Publishing Service.

Bechthold, C., & Bechthold, D. (1976). Signs for sounds. Hillsborough, CA: Educational Support Systems.

A P P E N D I X   D

Table 2

Treatment-Control Group Comparisons in the Database Grouped by Type of Phonics Program and Coded for Instructional Unit, Grade, Reading Ability of Participants, Length of Treatment, Type of Control Group, and Overall Effect size on Literacy Outcome Measures

Identity/Type[a] of Program	Inst. Unit	Grade/ Abil.[b]	Length[c]	Control[d]	d[e]
EMPHASIS ON SYNTHETIC PHONICS (S)					
74 Jolly Phonics (S)	Class	K at risk	12 wks.	Big Books (WL)	0.73
38 Successive phonics (S)	Sm gp	K at risk	8 wks.	Reg. curr.	0.62
03 Blachman PA (S)	Sm gp	K at risk	2-3 yrs.	Basal	0.72/0.36
04 SingSpellReadWrite (S)	Class	K	1 yr.	Basal	0.51
51 Lindamood PA (S)	Tutor	K at risk	3 yrs	Reg. curr.	0.33/0.67
72 Direct Instruction (S)	Class	K at risk	4 yrs.	Reg. curr.	-- /0.24
29 NRS-2 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.45
29 NRS-3 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.44
29 NRS-4 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.33
29 NRS-6 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.70
12 Synthetic basal (S)	Class	1st	1 yr	Whole word	2.27
04 SingSpellReadWrite (S)	Class	1st	1 yr.	Basal	0.25
15 Lippincott (S)	Class	1st	1 yr.	Whole word	0.84
48 Direct Instruction (S)	Sm gp	1st	1 yr.	Basal/Prev. yr.	0.38[f]
28 Direct Instruction (S)	Tutor	1st	10 wks.	Misc. (child reads)	1.99
08 Modif. Whole Lang (S)	Class	1st at risk	1 yr.	Basal	0.63
11 Open Court (S)	Class	1st at risk	1 yr.	Whole language	0.91
52 Direct Instruction (S)	Class	1st at risk	1 yr.	Whole language	0.07
69 Direct Instruction (S)	Sm gp	1st at risk	1 yr.	Basal	1.19
05 Lippincott (S)	Sm gp	1st at risk	2 yrs.	Whole word	0.48/.52
72 Direct Instruction (S)	Class	1st at risk	3 yrs.	Reg. curr.	--/0.00
04 SingSpellReadWrite (S)	Class	2nd	1 yr.	Basal	0.38
57 Sequential phonics (S)	Class	2nd	1 yr.	Whole language	-0.47
11 Open Court (S)	Class	2nd lo ach.	1 yr.	Whole language	0.12
37 Direct Instruction (S)	Class	gr 1-6 lo ach	10 wks.	Reg. curr.	0.01
55 Orton-Gillingham (S)	Class	3rd	1 yr.	Previous prog. (RC)	0.04
55 Orton-Gillingham (S)	Class	4th	1 yr.	Previous prog. (RC)	0.04
55 Orton-Gillingham (S)	Class	5th	1 yr.	Previous prog. (RC)	0.61
55 Orton-Gillingham (S)	Class	6th	1 yr.	Previous prog. (RC)	0.43
33 Lovett Dir. Inst. (S)	Sm gp	gr 2-3 RD	9 wks(35 s)	Misc. (Study skills)	0.24
33 Lovett Dir. Inst. (S)	Sm gp	gr 4 RD	9 wks(35 s)	Misc. (Study skills)	1.42
33 Lovett Dir. Inst. (S)	Sm gp	gr 5-6 RD	9 wks(35 s)	Misc. (Study skills)	0.09
17 Intersensory method (S)	Tutor	age 7-13 RD	18 wks.	Misc. (Subj. tutor)	0.53

Table 2 (Continued)

75 Lovett Dir. Inst. (S)	Sm gp	age 6-13 RD	70 hrs	Misc.(Study+Math)	0.24
32 Decoding skills (S)	Sm gp	age 8-13 RD	40 sessions	Misc. (Study skills)	0.39
47 Lippincott (S)	Sm gp	3rd RD	1 yr.	Whole word	0.50
13 Orton-Gillingham (S)	Sm gp	gr 2-3 RD	1 yr.	Whole word	0.27
41 Orton-Gillingham (S)	Sm gp	M=11yr RD	2 yrs.	Reg. curr.	/--0.54
47 Orton-Gillingham (S)	Sm gp	3rd RD	1 yr.	Whole word	0.04
55 Orton-Gillingham (S)	Class	3rd lo ach.	1 yr.	Previous prog. (RC)	0.63
55 Orton-Gillingham (S)	Class	4th lo ach.	1 yr.	Previous prog. (RC)	0.19
55 Orton-Gillingham (S)	Class	5th lo ach.	1 yr.	Previous prog. (RC)	-0.20
55 Orton-Gillingham (S)	Class	6th lo ach.	1 yr.	Previous prog. (RC)	0.13

Table 2 (Continued)

Identity/Type[a] of Program	Inst. Unit	Grade/ Abil.[b]	Length[c]	Control[d]	$\underline{d}$[e]
EMPHASIS ON BLENDING LARGER SUBUNITS AS WELL AS PHONEMES (LU)					
51 Embedded (LU)	Tutor	K at risk	3 yrs.	Reg. curr.	0.32/0.17
11 Embedded (LU)	Class	1st at risk	1 yr.	Whole language	0.36
11 Embedded (LU)	Class	2nd lo ach	1 yr.	Whole language	0.03
13 Onset-rime (LU)	Sm gp	gr 2-3 RD	1 yr.	Whole word	-0.11
44 RRD[g]-Early Steps (LU)	Tutor	1st at risk	1 yr.	Whole language	0.76
53 RRD[g]-Phonograms (LU)	Tutor	1st at risk	42 sessions	Reg. curr.	3.71
18 RRD[g]-Rime anal.(LU)	Tutor	gr 2-5 lo ach	11 wks.	Whole word	0.37
33 Lovett Analogy (LU)	Sm gp	gr 2/3 RD	9 wks (35 s)	Misc. (Study skills)	0.49
33 Lovett Analogy (LU)	Sm gp	gr 4 RD	9 wks (35 s)	Misc. (Study skills)	1.41
33 Lovett Analogy (LU)	Sm gp	gr 5/6 RD	9 wks (35 s)	Misc. (Study skills)	-0.25
75 Lovett Analogy (LU)	Sm gp	age 6-13 RD	70 hrs	Misc.(Study+Math)	0.50
COMBINATION PROGRAMS (C)					
75 Dir.Inst.+Analogy.(C)	Sm gp	age 6-13 RD	70 hrs	Misc.(Study+Math)	0.60
75 Analogy+Dir.Inst. (C)	Sm gp	age 6-13 RD	70 hrs	Misc.(Study+Math)	0.21
MISCELLANEOUS PHONICS (M)					
54 Developmental (M)	Sm gp	K at risk	12 wk.	Reg. curr. extended	0.47
09 Traditional basal (M)	Class	1st	1 yr.	Whole language	0.60
22 Analyze phonemes (M)	Sm gp	1st	6 wks.	Whole word	-0.07
22 Analyze onset-rimes (M)	Sm gp	1st	6 wks.	Whole word	0.14
26 Traditional basal (M)	Class	1st	1 yr.	Whole language	0.20
59 Sequential phonics (M)	Class	1st	1 yr/less	Whole language	0.00
60 Traditional basal (M)	Class	1st	1 yr.	Whole language	-0.33
36 Phonetic read/spell (M)	Tutor	1st at risk	1 yr (50 s)	Reg. curr.	0.53
35 Spelling mastery (M)	Class	2nd	1 yr.	Tradit. spell (RC)	0.38
34 Analytic (M)	Sm gp	age 7-13 RD	9 wks (35 s)	Misc. (Study skills)	0.16

[a] The programs listed as Direct Instruction include Reading Mastery and DISTAR.

[b] Information about grade/reading ability refers to the point in time when instruction began. RD refers to children classified as reading disabled. Lo ach refers to children above first grade who were identified as low achievers in their ability to read. At risk refers to kindergartners or first graders who performed poorly either on reading tests or on tests predictive of poor reading. If not marked, the sample consisted of normally developing readers.

[c] s refers to the number of sessions.

[d] RC means regular curriculum. WL means whole language. Misc. means miscellaneous category.

[e] Effect sizes listed singly are those observed at the end of training that lasted one year or less. When training lasted longer than one year, the first effect size reports the outcome at the end of the first year and the second effect size reports the outcome at the end of training.

[f] This effect size was not measured immediately after training but following a delay of six months.

[g] RRD refers to a program derived from Reading Recovery that was modified to include systematic phonics instruction in which phonemes were taught along with larger phonological units such as onsets, rimes and spelling patterns.

APPENDIX E

Table 3

Mean Effect Sizes (d) as a Function of Moderator Variables and Tests to Determine Whether Effect Sizes Were Significantly Greater Than Zero at $p < 0.05$, Whether Effect Sizes Were Homogeneous at $p < 0.05$, and Whether Effect Sizes Differed From Each Other at $p < 0.05$. Effect Sizes Refer to Outcomes Immediately After Training or At the End of One School Year, Whichever Came First, Unless Labeled as Followup or End of Training.

Moderator Variables and Levels	No. Cases	Mean d	Homogen.	95% CI	Contrasts
Time of Posttest					
End of Training	65	0.41*	No	0.36 to 0.47	n.s.
End of Training or One Year[a]	62	0.44*	No	0.38 to 0.50	
Followup	7	0.28*	Yes	0.10 to 0.46	
End of Training[b]	6	0.51*	Yes	0.32 to 0.70	n.s.
Followup	6	0.27*	Yes	0.07 to 0.46	
Outcome Measures					
Decoding regular words	30	0.67*	No	0.57 to 0.77	DecR = DecP;
Decoding pseudowords	40	0.60*	No	0.52 to 0.67	Both >
Reading misc. words	59	0.40*	No	0.34 to 0.46	RW, Spel,
Spelling words	37	0.35*	No	0.28 to 0.43	Oral,
Reading text orally	16	0.25*	No	0.15 to 0.36	Comp.
Comprehending text	35	0.27*	No	0.19 to 0.36	
Characteristics of Participants					
Grade					
Kind. & First	30	0.55*	No	0.47 to 0.62	K-1st >
2nd-6th, RD	32	0.27*	Yes	0.18 to 0.36	2nd-6th/RD
Younger Grades					
Kindergarten	7	0.56*	Yes	0.40 to 0.73	
First Grade	23	0.54*	No	0.46 to 0.63	
Kindergarten and First Graders on Outcome Measures					
Decoding regular words	8	0.98*	No	0.81 to 1.16	DecR >
Decoding pseudowords	14	0.67*	No	0.56 to 0.78	RMW, Co, Or;
Reading misc. words	23	0.45*	No	0.37 to 0.53	Spel > Or;
Spelling words	13	0.67*	No	0.54 to 0.79	DecP > Or.
Reading text orally	6	0.23*	No	0.05 to 0.41	
Comprehending text	11	0.51*	No	0.36 to 0.65	
2nd-6th, RD on Outcome Measures					
Decoding regular words	17	0.49*	No	0.34 to 0.65	DecR > Sp;
Decoding pseudowords	13	0.52*	Yes	0.37 to 0.66	DecP >
Reading misc. words	23	0.33*	No	0.22 to 0.44	Sp,Co.
Spelling words	13	0.09ns	Yes	-0.04 to 0.23	
Reading text orally	6	0.24*	Yes	0.08 to 0.39	
Comprehending text	11	0.12ns	Yes	-0.04 to 0.28	

Table 3 (Continued)

Moderator Variables and Levels	No. Cases	Mean d	Homogen.	95% CI	Contrasts
Grade and Reading Ability					
Kindergarten At Risk	6	0.58*	Yes	0.40 to 0.77	1AR >
1st Normal	14	0.48*	No	0.38 to 0.58	2N, 2AR,
1st At Risk	9	0.74*	No	0.56 to 0.91	RD
2nd-6th Normal	7	0.27*	Yes	0.12 to 0.43	
2nd-6th Lo Achievers	8	0.15ns	Yes	-0.06 to 0.36	
Reading Disabled	17	0.32*	Yes	0.18 to 0.46	
Socioeconomic Status					
Low SES	6	0.66*	Yes	0.48 to 0.85	n.s.
Middle SES	10	0.44*	No	0.28 to 0.60	
Varied	14	0.37*	Yes	0.26 to 0.48	
Not Given	32	0.43*	No	0.34 to 0.51	
Characteristics of Instruction					
Type of Phonics Program					
Synthetic	39	0.45*	No	0.39 to 0.52	n.s.
Larger Phon. Units	11	0.34*[d]	No	0.16 to 0.52	
Miscellaneous	10	0.27*	Yes	0.08 to 0.46	
Specific Phonics Programs					
NRS-Beck LRDC (S)	4	0.47*	Yes	0.33 to 0.60	n.s.
Direct Instruction (S)	4	0.48*	No	0.13 to 0.83	
Lovett Direct Instruct (S)	4	0.41*	Yes	0.04 to 0.77	
Lovett Analogy (LU)	4	0.48*	Yes	0.11 to 0.86	
Lippincott (S)	3	0.68*	Yes	0.43 to 0.93	
Orton Gillingham (S)	10	0.23*	Yes	0.06 to 0.39	
Sing Spell Read Write (S)	3	0.35*	Yes	0.21 to 0.50	
Synthetic Phonics For Various Readers Groups					
K & 1st At Risk[c]	9	0.64*	Yes	0.49 to 0.80	K&1AR >
1st Normal	8	0.54*	No	0.43 to 0.65	2-6LA,
2nd-6th Normal	6	0.27*	Yes	0.11 to 0.43	2-6N
2nd-6th Lo Achievers	6	0.14ns	Yes	-0.10 to 0.39	
Reading Disabled	9	0.36*	Yes	0.18 to 0.54	
Unit of Instruction					
Tutor	8	0.57*[d]	No	0.38 to 0.77	n.s.
Small Group	27	0.43*	Yes	0.34 to 0.52	
Class	27	0.39*	No	0.31 to 0.48	
Type of Control Group					
Basal	10	0.46*	Yes	0.37 to 0.55	n.s.
Regular Curriculum	16	0.41*	No	0.27 to 0.54	
Whole Language	12	0.31*	No	0.16 to 0.47	
Whole Word	10	0.51*	No	0.35 to 0.67	
Miscellaneous	14	0.46*	Yes	0.28 to 0.63	

Table 3 (Continued)

Moderator Variables and Levels	No. Cases	Mean $\underline{d}$	Homogen	95% CI	Contrasts
Characteristics of the Design of Studies					
Assignment of Participants to Treatment and Control Groups					
Random	23	0.45*	Yes	0.32 to 0.58	n.s.
Nonequivalent Groups	39	0.43*	No	0.37 to 0.50	
Sample Size					
20 to 31	14	0.48*	No	0.26 to 0.70	n.s.
32 to 52	16	0.31*	Yes	0.15 to 0.47	
53 to 79	16	0.36*	No	0.23 to 0.49	
80 to 320	16	0.49*	No	0.41 to 0.57	

* indicates that effect size was significantly greater than zero at $\underline{p} < 0.05$.

  ns indicates not significantly different from zero.

[a] Effect sizes indicate literacy outcomes at the end of training for studies lasting 1 year or less, and at the end of the first school year for studies that continued training beyond 1 year.

[b] The six studies in both comparisons were the same studies.

[c] The kindergarten and 1st grade at-risk groups had identical $\underline{d}$s and were combined.

[d] This effect size was adjusted to reduce the impact of one atypically large outlier.

Table 4

Characteristics of Sets of Studies of Special Interest

Type of Program[a]	Inst. Unit	Grade/ Abil.	Length	Control	_d_
STUDIES WITH TRAINING LASTING MORE THAN A YEAR[c]					
03 Blachman PA (S)	Sm gp	K at risk	2-3 yrs	Basal	0.72/0.64/0.36
51 Lindamood PA (S)	Tutor	K at risk	3 yrs	Reg. curr.	0.33/0.75/0.67
51 Embedded (LU)	Tutor	K at risk	3 yrs	Reg. curr.	0.32/0.28/0.17
05 Lippincott (S)	Sm gp	1st at risk	2 yrs	Whole word	0.48/0.52
72 Direct Instruction (S)	Class	K at risk	4 yrs	Reg. curr.	--/0.24
72 Direct Instruction (S)	Class	1st at risk	3 yrs	Reg. curr.	--/0.00
41 Orton-Gillingham (S)	Sm gp	M=11yr RD	2 yrs	Reg. curr.	--/0.54
STUDIES MEASURING IMMEDIATE OUTCOMES AND LONG-TERM OUTCOMES[b]					
18 Rime analogy (LU)	Tutor	gr 2-5 lo ach	11 wks	Whole word	0.37/0.56 (1 yr.)
36 Phonetic read/spell (M)	Tutor	1st at risk	50 ses	Reg. curr.	0.53/0.32 (1 yr.)
44 Early Steps (LU)	Tutor	1st at risk	1 yr	Whole language	0.76/0.86 (4 mo.)
47 Orton-Gillingham (S)	Sm gp	3rd RD	1 yr	Whole word	0.04/-0.47 (6 mo.)
47 Lippincott (S)	Sm gp	3rd RD	1 yr	Whole word	0.50/0.33 (6 mo.)
48 Direct Instruction (S)	Sm gp	1st	1 yr	Basal (Prev. yr)	--/0.38 (6 mo.)
74 Jolly Phonics (S)	Class	K at risk	12 wks	Big Books (WL)	0.73/0.28 (1 yr.)
2ND-6TH LOW ACHIEVERS					
11 Embedded (LU)	Class	2nd lo ach	1 yr	Whole language	0.03
11 Open Court (S)	Class	2nd lo ach.	1 yr	Whole language	0.12
18 Rime analogy (LU)	Tutor	gr 2-5 lo ach	11 wks	Whole word	0.37
37 Direct Instruction (S)	Class	gr 1-6 lo ach	10 wks	Reg. curr.	0.01
55 Orton-Gillingham (S)	Class	3rd lo ach.	1 yr	Previous prog. (RC)	0.63
55 Orton-Gillingham (S)	Class	5th lo ach.	1 yr	Previous prog. (RC)	-0.20
55 Orton-Gillingham (S)	Class	6th lo ach.	1 yr	Previous prog. (RC)	0.13
55 Orton-Gillingham (S)	Class	4th lo ach.	1 yr	Previous prog. (RC)	0.19
TUTORING COMPARISONS					
51 Lindamood PA (S)	Tutor	K at risk	3 yrs	Reg. curr. (class)	0.33/0.67
51 Embedded (LU)	Tutor	K at risk	3 yrs	Reg. curr. (class)	0.32/0.17
28 Direct Instruction (S)	Tutor	1st	10 wks	Misc. (child reads) (tutor)	1.99
36 Phonetic read/spell (M)	Tutor	1st at risk	50 ses	Reg. curr. (class)	0.53
44 Early Steps (LU)	Tutor	1st at risk	1 yr	Whole lang. (sm gp)	0.76
53 Phonograms (LU)	Tutor	1st at risk	42 ses	Reg. curr. (class)	3.71
17 Intersensory method (S)	Tutor	age 7-13 RD	18 wks	Misc. (Subj. tutor)	0.53
18 Rime analogy (LU)	Tutor	gr 2-5 lo ach	11 wks	Whole word (tutor)	0.37

_____

[a] Letters in parentheses refer to the type of phonics program: S (synthetic), LU (Larger subunits), M (Miscellaneous).

[b] The first effect size is for the immediate posttest and the second is for the delayed posttest. The length of the delay between posttests is given in parentheses.

[c] When 3 effect sizes are reported, these refer to effects at the end of each year of training.

Table 5

Number of Comparisons by Grade and Reading Ability

Grade	Reading Ability			
	Normally Developing	At Risk/ Low Achievers	Reading Disabled	Total
Kindergarten	1	6 (K-AR)	--	7
First Grade	14 (1N)	9 (1-AR)	--	23
Second Grade	3 (2-6N)	2 (2-6 AR)	--	5
3rd-6th Grades	4 (2-6N)	4 (2-6 AR)	6 (RD)	14
Mixed grades	--	2 (2-6 AR)	11 (RD)	13
Total	22	23	17	62

Note. The symbols in parentheses refer to the groups that were created for the meta-analysis.

A p p e n d i x   F

Table 6

Characteristics of the Treatment-Control Group Comparisons Utilizing Specific Phonics Programs That Were Included in the Meta-Analysis

Identify/Type of Program	Inst. Unit	Grade/ Abil.	Length	Control	d
28 Direct Instruction (S)	Tutor	1st	10 wks.	Misc. (child reads)	1.99
52 Direct Instruction (S)	Class	1st at risk	1 yr.	Whole language	0.07
69 Direct Instruction (S)	Sm gp	1st at risk	1 yr.	Basal	1.19
37 Direct Instruction (S)	Class	gr 1-6 lo ach	10 wks.	Reg. curr.	0.01
33 Lovett Dir. Inst. (S)	Sm gp	gr 4 RD	9 wks(35 s)	Misc. (Study skills)	1.42
33 Lovett Dir. Inst. (S)	Sm gp	gr 2-3 RD	9 wks(35 s)	Misc. (Study skills)	0.24
33 Lovett Dir. Inst. (S)	Sm gp	gr 5-6 RD	9 wks(35 s)	Misc. (Study skills)	0.09
75 Lovett Dir. Inst. (S)	Sm gp	age 6-13 RD	70 hrs	Misc.(Study+Math)	0.24
33 Lovett Analogy (LU)	Sm gp	gr 4 RD	9 wks (35 s)	Misc. (Study skills)	1.41
33 Lovett Analogy (LU)	Sm gp	gr 2/3 RD	9 wks (35 s)	Misc. (Study skills)	0.49
33 Lovett Analogy (LU)	Sm gp	gr 5/6 RD	9 wks (35 s)	Misc. (Study skills)	-0.25
75 Lovett Analogy (LU)	Sm gp	age 6-13 RD	70 hrs	Misc.(Study+Math)	0.50
15 Lippincott (S)	Class	1st	1 yr.	Whole word	0.84
05 Lippincott (S)	Sm gp	1st at risk	2 yrs.	Whole word	0.48
47 Lippincott (S)	Sm gp	3rd RD	1 yr.	Whole word	0.50
29 NRS-6 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.70
29 NRS-4 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.33
29 NRS-3 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.44
29 NRS-2 (Beck) (S)	Sm gp	1st	1 yr.	Basal	0.45
55 Orton-Gillingham (S)	Class	3rd	1 yr.	Previous prog. (RC)	0.04
55 Orton-Gillingham (S)	Class	4th	1 yr.	Previous prog. (RC)	0.04
55 Orton-Gillingham (S)	Class	5th	1 yr.	Previous prog. (RC)	0.61
55 Orton-Gillingham (S)	Class	6th	1 yr.	Previous prog. (RC)	0.43
55 Orton-Gillingham (S)	Class	3rd lo ach.	1 yr.	Previous prog. (RC)	0.63
55 Orton-Gillingham (S)	Class	4th lo ach.	1 yr.	Previous prog. (RC)	0.19
55 Orton-Gillingham (S)	Class	5th lo ach.	1 yr.	Previous prog. (RC)	-0.20
55 Orton-Gillingham (S)	Class	6th lo ach.	1 yr.	Previous prog. (RC)	0.13
13 Orton-Gillingham (S)	Sm gp	gr 2-3 RD	1 yr.	Whole word	0.27
47 Orton-Gillingham (S)	Sm gp	3rd RD	1 yr.	Whole word	0.04
04 SingSpellReadWrite (S)	Class	K	1 yr.	Basal	0.51
04 SingSpellReadWrite (S)	Class	1st	1 yr.	Basal	0.25
04 SingSpellReadWrite (S)	Class	2nd	1 yr.	Basal	0.38

Appendix F (continued)

Table 7

Descriptions of the Specific Phonics Programs Examined in the Meta-Analysis

1. <u>Direct Instruction</u>. The Direct Instruction program is based on a behavioral analysis of the steps involved in learning to decode (Carnine & Silbert, 1979; Engelmann, 1980; Engelmann & Bruner, 1969, 1978, 1988; Engelmann & Osborn, 1987; Kameenui et al., 1997). At the beginning of the program, students are not taught letter names but only letter-sound relations through highly structured instruction that uses cueing and reinforcement procedures derived from a behavioral analyses of instruction. The task of decoding is broken down into its component parts, and each of these parts is taught separately, from letter sounds to blending to reading words in context. Instruction is scripted and the lessons are fast paced, with high student participation. The text for the first-year program is written in a script that, although it preserves English spelling, contains printed marks that cue the reader about silent letters and different vowel sounds. Children practice in specially constructed books containing taught sounds, although children may be encouraged to read widely in children's literature as well (e.g., Meyer, 1983).

2. <u>Lovett Direct Instruction</u>. The synthetic phonics program used by Lovett and Steinbach (1997) and Lovett et al. (in press) adopts the Direction Instruction model to remediate the decoding and phonemic awareness difficulties of severely disabled readers. Children are taught phonological analysis and blending (phonemic awareness) orally and also letter sound associations in the context of word recognition and decoding instruction. The program focuses on training sound blending and acquisition of a left-to-right phonological decoding strategy. The special orthography highlights salient features of many letters and provides visual cues such as symbols over long vowels, letter size variations, and connected letters to facilitate learning. Cumulative, systematic review and many opportunities for overlearning are hallmarks of this approach. New material is not introduced until the child fully masters previously instructed material.

3. <u>Lovett Analogy</u>. A second program also used with severely disabled readers by Lovett and Steinbach (1997) and Lovett et al. (in press) was adapted from the Benchmark Word Identification/Vocabulary Development program developed by Gaskins et al. (1986). This program is strongly metacognitive in its focus. It teaches children how to use four metacognitive strategies to decode words: reading words by analogy, detecting parts of words that are known, varying the pronunciations of vowels to maintain flexibility in decoding attempts, and "peeling off" prefixes and suffixes in words. Children learn a set of 120 key words exemplifying high-frequency spelling patterns, 5 words per day. They learn to segment the words into subunits so that they can use these known words and their parts to read other similarly spelled words. They learn letter-sound associations for vowels and affixes. Various types of texts provide children with practice applying the strategies taught.

4. <u>Lippincott</u>. The Lippincott Basic Reading Series (McCracken & Walcutt, 1963, 1975) is a direct code method which, from the outset, approaches reading from a phonic/linguistic perspective. Beginning with children's spoken language, the Lippincott program teaches in a systematic manner how to use the alphabetic code to move from printed words to oral language. Instruction begins with short-a and builds knowledge of regular sound/symbol relationships. Children are first taught to decode phonetically regular words, with blending of phonic elements directly taught. Once they are proficient, long vowels and irregular spellings are introduced. Although the primary instructional focus is on decoding, another goal of this method is the instant recognition of words. However, rather than relying on a "context clue" approach to word recognition, children are taught how and why the letters come to

Appendix F (continued)

Table 7 (continued)

represent these words, and they learn to "break the code" to decipher new words independently. Review and reinforcement are an integral part of the program. Spelling is sometimes taught as one component of the reading lesson with spelling lists developed from the words introduced in each unit of reading instruction (Brown & Felton, 1990).

5. <u>NRS by Beck and Mitroff</u>. The New Primary Grades Reading System for an Individualized Classroom (NRS) was developed by Beck and Mitroff (1972). It is a code-breaking approach. The program begins by teaching self-management skills, letter-sound correspondences, and chain blending to decode words. Children are taught to pronounce the first letter of a word followed by the second letter and then to blend the two sounds; then they pronounce the third letter and add it to the blend. In the first lesson, children are taught five isolated letter-sound relations, and once they are known, children are immediately taught to blend them to form real words. Subsequent letter-sounds are taught one at a time and blended with the earlier letters. Not only synthetic phonics but also analytic phonics is taught as children explore words and their parts. The method is linguistic as well because the major spelling patterns of words are displayed in texts to draw attention to similarities and contrasts, and because there is minimum teaching of explicit pronunciation rules. Instruction is individualized. After the first two levels, children work through the curriculum at different rates.

6. <u>Orton Gillingham</u>. The Orton-Gillingham approach (Cox, 1991; Gillingham & Stillman, 1979) begins with the direct teaching of individual letters paired with their sounds using a Visual-Auditory-Kinesthetic-Tactile (VAKT) procedure that involves tracing the letter while saying its name and sound, blending letters together to read words and sentences, and finally reading short stories constructed to contain only taught sounds. Spelling words from dictation is also part of an Orton-Gillingham lesson. Each letter-sound is learned to mastery through repetition. More advanced lessons involve teaching learners to blend syllables together and read more complex texts. Among those approaches based on Orton and Gillingham's work are the Slingerland approach (Lovitt & DeMier, 1984), the Spaulding Approach, Recipe for Reading, and Alphabetic Phonics (Ogden, Hindman, & Turner, 1989). There are differences among these approaches, largely in the sequencing of materials, but they all have the general characteristics discussed.

7. <u>Sing, Spell, Read & Write</u>. The Sing, Spell, Read and Write (SSRW) program (Dickson, 1972) also teaches synthetic phonics. It consists of several charts, books (both readers and workbooks), letter and word cards, tests, and audio tapes. The tapes contain songs about several phonics generalizations. Through the tapes, the students learn the sounds of letters and letter combinations. Also songs combined with charts help students learn the spellings of words. The lessons begin by teaching letter-sounds in isolation for each letter of the alphabet. When students have mastered certain sounds, they begin reading phonetic storybooks. The first five books each focus on a different vowel sound. The remaining books expand the vocabulary in a way that is consistent with the letter-sounds taught. Students are taught to spell the words they learn to read, with the words presented in sentences. Most of the writing students do involves filling in blanks or answering questions related to words being learned. The program has a "racetrack" which is posted in classrooms and notes students' progress by placement of a race car on the chart (Bond et al., 1995-96).

Appendix G

Studies in the Phonics Database, Their Characteristics, and Effect Sizes

(*Note:* key to this chart is on page 2-176)

Author and Year, Treatment	Characteristics of Training				Characteristics of Part.			Features of Design				Effect Sizes on Post-tests							
	Type of Phonics	Control Group	Tr.unit	Length of Training	Grade/ Age	Reading Ability	SES	Group Assign.	Sig Pre-test Diff	Total N	Time of Post-test	Mean	Word ID	Dec	Spell	Comp	Nonw	Oral Read	Gen. Read
03 - Blachman et al., (1999)	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Blachman PA	Syn	Basal	SmG	2-3yrs(41s-,20m/d)	K	AR	Low	NE	No	159	Imm.	0.72	-0.17	1.08	0.94	.	1.04	.	.
Blachman PA				(1st gr=30 m/d)						128	2nd yr tr.	0.64	0.35	0.81	0.53	.	0.86	.	.
Blachman PA				(2nd gr=30 m/d)						106	3rd yr tr.	0.36	0.42	0.55	0	.	0.45	.	.
04 - Bond et al., 1995																			
Sing, Spell, Read, Write	Syn	Basal	Class	1 yr.(20 lessons)	K	N	Var	NE	No	144	Imm.	0.51	0.38	.	.	.	1.01	0.13	.
Sing, Spell, Read, Write	Syn	Basal	Class	1 yr.	1st	N	Var	NE	No	276	Imm.	0.25	0.23	.	0.14	.	0.6	0.03	.
Sing, Spell, Read, Write	Syn	Basal	Class	1 yr.	2nd	N	Var	NE	No	320	Imm.	0.38	0.44	.	0.18	.	0.55	0.33	.
05 - Brown & Felton, 1990																			
Lippincott	Syn	Wh.W.	SmG	2 yrs.	1st	AR	NG	R	No	47	Imm.	0.48	0.02	.	0.51	.	0.92	.	.
Lippincott											2nd yr tr.	0.52	0.51	0.63	0.38	.	0.55	.	.
08 - Eldredge, 1991																			
Modified Whole Language	Syn	Basal	Class	1 yr. (15m/d)	1st	AR	Low	NE	No	105	Imm.	0.63	.	.	.	0.83	0.43	.	.
09 - Evans, 1985																			
Traditional Basal	Misc	Wh.L.	Class	1 yr.	1st	N	Var	NE	NG	20*(N=-247)	Imm.	0.6	.	.	.	0.6	.	.	.

Appendix G (continued)

Author and Year, Treatment	Characteristics of Training				Characteristics of Part.			Features of Design				Effect Sizes on Post-tests							
	Type of Phonics	Control Group	Tr.unit	Length of Training	Grade/ Age	Reading Ability	SES	Group Assign.	Sig Pre-test Diff	Total N	Time of Post-test	Mean	Word ID	Dec	Spell	Comp	Nonw	Oral Read	Gen. Read
11 - Foorman et al., 1998																			
Open Court	Syn	Wh.L.	Class	1 yr. (30m/d)	1st	AR	Var	NE	NG	68	Imm.	0.91	1.63	1.14	0.56	0.32	.	.	.
Embedded	LU	Wh.L.	Class	1 yr.	1st	AR	Var	NE	NG	70	Imm.	0.36	0.56	0.51	0.26	0.1	.	.	.
Open Court	Syn	Wh.L.	Class	1 yr.	2nd	LA	Var	NE	NG	35	Imm.	0.12	0.52	0.32	-0.19	-0.19	.	.	.
Embedded	LU	Wh.L.	Class	1 yr.	2nd	LA	Var	NE	NG	57	Imm.	0.03	0.37	0.22	-0.25	-0.24	.	.	.
12 - Foorman et al., 1991																			
Synthetic basal	Syn	Wh.W.	Class	1 yr. (45 m/d)	1st	N	Mid	NE	No	6*(N=8-0)	Imm.	2.27	1.92	2.67	2.21	.	.	.	.
13 - Foorman et al., 1997																			
Orton-Gillingham	Syn	Wh.W.	SmG	1 yr. (60 m/d)	gr 2-3	RD	Mid	NG	Yes	67	Imm.	0.27	0.17	0.58	0.05	.	.	.	.
Onset-rime	LU	Wh.W.	SmG	1 yr.	gr 2-3	RD	Mid	NG	Yes	85	Imm.	-0.11	-0.19	0.09	-0.23	.	.	.	.
15 - Fulwiler & Groff, 1980																			
Lippincott	Syn	Wh.W.	Class	1 yr.	1st	N	NG	NE	NG	147	Imm.	0.84	.	0.91	.	0.76	.	.	.
17 - Gittelman & Feingold, 1983																			
Intersensory Method	Syn	Misc.	Tutor	18 wks.(54s)	7-13yr	RD	Mid	R	No	56	Imm.	0.53	0.76	0.67	0.12	0.57	.	.	.
18 - Greaney et al., 1997																			
RRD-Rime analogy	LU	Wh.W.	Tutor	11 wks(31s,3-0m)	gr 2-5	LA	NG	R	No	36	Imm.	0.37	0.39	.	.	.	0.51	0.2	.
RRD-Rime analogy										34	follow up	0.56	0.47	.	.	.	0.76	0.44	.

Author and Year, Treatment	Characteristics of Training				Characteristics of Part.			Features of Design				Effect Sizes on Post-tests							
	Type of Phonics	Control Group	Tr.unit	Length of Training	Grade/ Age	Reading Ability	SES	Group Assign.	Sig Pre-test Diff	Total N	Time of Post-test	Mean	Word ID	Dec	Spell	Comp	Nonw	Oral Read	Gen. Read
22 - Haskell et al., 1992																			
Analyze Onset-Rimes	Misc	Wh.W.	SmG	6 wks(15s, 20m)	1st	N	Mid	R	No	24	Imm.	0.14	0.2	0.09	.	.	.	.	.
Analyze Phonemes	Misc	Wh.W.	SmG	6 wks(15s, 20m)	1st	N	Mid	R	No	24	Imm.	-0.07	-0.08	-0.06	.	.	.	.	.
26 - Klesius et al., 1991																			
Traditional Basal	Misc	Wh.L.	Class	1 yr.	1st	N	Var	NE	Yes	6*(N=1-12)	Imm.	0.2	.	.	0.36	0.18	0.07	.	.
28 - Leach & Siddall, 1990																			
Direct Instruction	Syn	Misc.	Tutor	10 wks (15m/d)	1st	N	NG	R	No	20	Imm.	1.99	.	.	.	1.8	.	2.18	.
29 - Leinhardt & Engel, 1981																			
NRS-study 2 (Beck)	Syn	Basal	SmG	1 yr.	1st	N	NG	NE	Yes	187	Imm.	0.45	0.45	.	.	.	.	.	.
NRS-study 3 (Beck)	Syn	Basal	SmG	1 yr.	1st	N	NG	NE	Yes	263	Imm.	0.44	0.44	.	.	.	.	.	.
NRS-study 4 (Beck)	Syn	Basal	SmG	1 yr.	1st	N	NG	NE	Yes	256	Imm.	0.33	0.33	.	.	.	.	.	.
NRS-study 6 (Beck)	Syn	Basal	SmG	1 yr.	1st	N	NG	NE	Yes	241	Imm.	0.7	0.7	.	.	.	.	.	.
32 - Lovett et al., 1989																			
Decoding Skills	Syn	Misc.	SmG	40 ses (33-40h)	8-13yr	RD	Mid	R	No	118	Imm.	0.39	0.78	0.7	0.42	0.07	0.1	0.27	.
33 - Lovett & Steinbach, 1997																			
Lovett Analogy	LU	Misc.	SmG	9wks (35h)	gr 2/3	RD	NG	R	No	28	Imm.	0.49	-0.12	0.85	.	.	0.75	.	.
Lovett Analogy	LU	Misc.	SmG	9wks (35h)	gr 4	RD	NG	R	No	22	Imm.	1.41	0.84	2.06	.	.	1.33	.	.

Appendix G (continued)

Author and Year, Treatment	Characteristics of Training				Characteristics of Part.			Features of Design				Effect Sizes on Post-tests							
	Type of Phonics	Control Group	Tr.unit	Length of Training	Grade/ Age	Reading Ability	SES	Group Assign.	Sig Pre-test Diff	Total N	Time of Post-test	Mean	Word ID	Dec	Spell	Comp	Nonw	Oral Read	Gen. Read
Lovett Analogy	LU	Misc.	SmG	9wks (35h)	gr 5/6	RD	NG	R	No	24	Imm.	-0.25	-0.49	-0.15	.	.	-0.1	.	.
Lovett Direct Instruction	Syn	Misc.	SmG	9wks (35h)	gr 2/3	RD	NG	R	No	32	Imm.	0.24	0.02	0.24	.	.	0.46	.	.
Lovett Direct Instruction	Syn	Misc.	SmG	9wks (35h)	gr 4	RD	NG	R	No	25	Imm.	1.42	1.03	1.53	.	.	1.7	.	.
Lovett Direct Instruction	Syn	Misc.	SmG	9wks (35h)	gr 5/6	RD	NG	R	No	27	Imm.	0.09	-0.24	0.25	.	.	0.25	.	.
34 - Lovett et al., 1990																			
Analytic	Misc	Misc.	SmG	9wks (35h)	7-13yr	RD	Mid	R	NG	36	Imm.	0.16	0.13	0.11	0.23	.	.	.	.
35 - Lum & Morton, 1984																			
Spelling Mastery	Misc	Rg.cls.	Class	1 yr.(20-30 m/d)	2nd	N	NG	NE	No	36	Imm.	0.38	0.31	.	0.45	.	.	.	.
36 - Mantzicopoulos et al., 1992																			
Phonetic read/spell	Misc	Rg.cls.	Tutor	50s (1h/wk)	1st	AR	Mid	R	No	112	Imm.	0.53	.	.	.	.	0.53	.	.
Phonetic read/spell										112	follow up	0.32	.	0.33	0.3	0.08	0.56	.	.
37 - Marston et al., 1995																			
Direct Instruction	Syn	Rg.cls.	Class	10 wks (45m/d)	gr 1-6	LA	NG	NE	Y/Adj	53	Imm.	0.01	.	.	.	.	.	0.01	.
38 - Martinussen & Kirby, 1998																			
Successive phonics	Syn	Rg.cls.	SmG	8 wks(40-60m/wk)	K	AR	NG	R	No	26	Imm.	0.62	0.53	0.63	0.68	.	0.62	.	.
41 - Oakland et al., 1998																			
Orton-Gillingham	Syn	Rg.cls.	SmG	2 yrs.(350h)	M=11y	RD	NG	NE	Yes	48	2nd yr tr.	0.54	0.71	.	0.23	0.62	0.61	.	.

Author and Year, Treatment	Characteristics of Training				Characteristics of Part.			Features of Design				Effect Sizes on Post-tests							
	Type of Phonics	Control Group	Tr.unit	Length of Training	Grade/Age	Reading Ability	SES	Group Assign.	Sig Pre-test Diff	Total N	Time of Post-test	Mean	Word ID	Dec	Spell	Comp	Nonw	Oral Read	Gen. Read
44 - Santa & Hoien, 1999																			
RRD-Early Steps	LU	Wh.L.	Tutor	1 yr.(30m/d)	1st	AR	Var	NE	No	49	Imm.	0.76	0.93	.	0.63	0.73	.	.	.
RRD-Early Steps										41	follow up	0.86	0.57	.	.	0.87	1.15	.	.
47 - Silberberg et al., 1973																			
Lippincott	Syn	Wh.W.	SmG	1 yr.	gr 3	RD	NG	NE	Yes	69	Imm.	0.5	0.7	.	.	0.36	.	0.45	.
Orton-Gillingham	Syn	Wh.W.	SmG	1 yr.	gr 3	RD	NG	NE	Yes	65	Imm.	0.04	0.31	.	.	0.09	.	-0.29	.
Lippincott										62	follow up	0.33	0.37	.	.	-0.04	.	0.66	.
Orton-Gillingham										58	follow up	-0.47	-0.19	.	.	-0.81	.	-0.4	.
48 - Snider, 1990																			
Direct Instruction	Syn	Basal	SmG	1yr.(60m/d)	1st	N	Mid	NE	No	66	follow up	0.38	.	0.6	0.44	0.1	.	.	.
51 - Torgesen et al., 1999												.	.	.	.	.	.	.	.
Lindamood PA	Syn	Rg.cls.	Tutor	2.5 yrs.(80m/-wk)	K	AR	NG	R	No	65	Imm.	0.33	0.08	.	.	.	0.58	.	.
Embedded	LU	Rg.cls.	Tutor	2.5 yrs.(80m/-wk)	K	AR	NG	R	No	68	Imm.	0.32	0.52	.	.	.	0.12	.	.
Lindamood PA										65	2nd yr tr.	0.75	0.64	.	.	0.49	1.13	.	.
Embedded										68	2nd yr tr.	0.28	0.24	.	.	0.29	0.31	.	.
Lindamood PA										65	3rd yr tr.	0.67	0.67	.	0.64	0.36	1.01	.	.
Embedded										68	3rd yr tr.	0.17	0.25	.	0.1	0.17	0.16	.	.

Appendix G (continued)

Author and Year, Treatment	Characteristics of Training				Characteristics of Part.			Features of Design				Effect Sizes on Post-tests							
	Type of Phonics	Control Group	Tr.unit	Length of Training	Grade/ Age	Reading Ability	SES	Group Assign.	Sig Pre-test Diff	Total N	Time of Post-test	Mean	Word ID	Dec	Spell	Comp	Nonw	Oral Read	Gen. Read
52 - Traweek & Berninger, 1997																			
Direct Instruction	Syn	Wh.L.	Class	1yr.	1st	AR	Low	NE	Y/Adj	38	Imm.	0.07	0.07	.	.	.	.	.	.
53 - Tunmer & Hoover, 1993																			
RRD-Phonograms	LU	Rg.cls.	Tutor	42 s (30m/d)	1st	AR	NG	NG	NG	64	Imm.	3.71	2.94	.	1.63	.	1.49	8.79	.
54 - Vandervelden & Siegel, 1997																			
Developmental	Misc	Rg.cls.	SmG	12wks(30--45m/wk)	K	AR	Low	NE	No	29	Imm.	0.47	0.04	.	1.11	.	0.57	0.15	.
55 - Vickery et al., 1987																			
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	3rd	N	NG	NE	NG	63	Imm.	0.04	.	.	.	.	.	.	0.04
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	4th	N	NG	NE	NG	71	Imm.	0.04	.	.	.	.	.	.	0.04
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	5th	N	NG	NE	NG	74	Imm.	0.61	.	.	.	.	.	.	0.61
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	6th	N	NG	NE	NG	79	Imm.	0.43	.	.	.	.	.	.	0.43
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	3rd	LA	NG	NE	NG	46	Imm.	0.63	.	.	.	.	.	.	0.63
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	4th	LA	NG	NE	NG	47	Imm.	0.19	.	.	.	.	.	.	0.19
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	5th	LA	NG	NE	NG	45	Imm.	-0.2	.	.	.	.	.	.	-0.2
Orton-Gillingham	Syn	Rg.cls.	Class	1 yr.(55 m/d)	6th	LA	NG	NE	NG	41	Imm.	0.13	.	.	.	.	.	.	0.13
57 - Wilson & Norman, 1998																			
Sequential phonics	Syn	Wh.L.	Class	1 yr.	2nd	N	NG	NE	No	54	Imm.	-0.47	-0.33	.	.	-0.61	.	.	.

Author and Year, Treatment	Characteristics of Training				Characteristics of Part.			Features of Design				Effect Sizes on Post-tests							
	Type of Phonics	Control Group	Tr.unit	Length of Training	Grade/ Age	Reading Ability	SES	Group Assign.	Sig Pre-test Diff	Total N	Time of Post-test	Mean	Word ID	Dec	Spell	Comp	Nonw	Oral Read	Gen. Read
59 - Freppon, 1991																			
Sequential phonics	Misc	Wh.L.	Class	1 yr.	1st	N	Mid	NE	Yes	24	Imm.	0	.	.	.	.	.	0	.
60 - Griffith et al., 1992																			
Traditional basal	Misc	Wh.L.	Class	1 yr.	1st	N	NG	NE	No	24	Imm.	-0.33	-1.11	.	-0.54	-0.43	0.78	.	.
69 - Umbach et al., 1989																			
Direct Instruction	Syn	Basal	SmG	1 yr.(50 m/d)	1st	AR	Low	R	No	31	Imm.	1.19	1.3	.	.	1.08	.	.	.
72 - Gersten et al., 1988																			
Direct Instruction	Syn	Rg.cls.	Class	4 yrs.	K	AR	Low	NE	No	101	4th yr tr	0.24	.	.	0.16	0.28	.	.	0.27
Direct Instruction	Syn	Rg.cls.	Class	3 yrs.	1st	AR	Low	NE	No	141	3rd yr tr.	0	.	.	-0.12	0.11	.	.	0.02
74 - Stuart, 1999																			
Jolly Phonics	Syn	Wh.L.	Class	12 wks(60m/-d)	K	AR	Low	NE	Y/Adj	112	Imm.	0.73	0.56	.	1.11	0.36	0.9	.	.
Jolly Phonics										112	follow up	0.28	0.11	.	0.5	0.31	-0.03	0.49	.
75 - Lovett et al., (in press)																			
Dir. Instruction + Analogy	Com	Misc.	SmG	70h	6-13yr	RD	Var	R	NG	37	Imm.	0.6	0.36	1	0.15	0.27	1.22	.	.
Analogy + Direct Instruction	Com	Misc.	SmG	70h	6-13yr	RD	Var	R	NG	32	Imm.	0.21	0.04	0.55	-0.2	0.12	0.52	.	.
Lovett Direct Instruction	Syn	Misc.	SmG	70h	6-13yr	RD	Var	R	NG	40	Imm.	0.24	0.21	0.36	-0.19	0.42	0.42	.	.
Lovett Analogy	LU	Misc.	SmG	70h	6-13yr	RD	Var	R	NG	42	Imm.	0.5	0.47	0.75	0.01	0.6	0.66	.	.

A p p e n d i x   G   (continued)

Abbreviations Key

Following is a key to Appendix G.

Word ID = Word Identification	h = hour
Dec = Decoding	s = session(s)
Spell = Spelling	wks = weeks
Comp = Comprehension	gr = grade
Nonw = Nonword reading	M = mean
Oral Read = Oral reading	K = Kindergarten
Gen. Read = Generic reading	RD = Reading Disabled
Syn = Synthetic;	AR = At Risk
LU = Larger Units	LA = Low Achievement
Misc = Miscellaneous	NG = Not Given
Com = Combination	Var = Varied
Wh.W. = Whole Word	Mid = Middle class
Wh.L. = Whole Language	R = Random assignment
Rg. Cls. = Regular class	NE = Non Equivalent groups
SmG = Small group	Y/Adj = Yes, but means were adjusted for pretest differences
yr, = year	Imm. = Immediate
m = minutes	tr = training
m/d = minutes a day	*class was used as the unit of analysis