The Federal Government Wants Me to Teach What?
A Teacher's Guide to the National Reading Panel Report

Diane Stephens
University of South Carolina

National Council of Teachers of English
1111 W. Kenyon Road, Urbana, Illinois 61801-1096

Manuscript Editor: Jane Curran

Staff Editor: Kurt Austin

It is the policy of NCTE in its journals and other publications to provide a forum for the open discussion of ideas concerning the content and the teaching of English and the language arts. Publicity accorded to any particular point of view does not imply endorsement by the Executive Committee, the Board of Directors, or the membership at large, except in announcements of policy, where such endorsement is clearly specified.

Every effort has been made to provide current URLs and email addresses, but because of the rapidly changing nature of the Web, some sites and addresses may no longer be accessible.

Contents

Acknowledgments

Over a two-year period, a group of almost 30 teacher-educators in South Carolina read the National Reading Panel (NRP) report in full, and then each read a number of the studies included in the NRP report. Together we read the entire NRP report, all the articles cited by the Alphabetics, Fluency, and Vocabulary subgroups that we could find, and more than a hundred of the studies included in the Comprehension report. I personally read the entire NRP report, 78 of the 90 studies cited by the Alphabetics group, 10 of the 16 articles cited by the Fluency subgroup, and more than 150 of the 250 articles cited by the Comprehension subgroup. For this last subgroup, I narrowed my review to the studies in the categories the subgroup members considered effective, eliminating only those that were dissertations, speeches, or written in Spanish.

My discussions with these educators about the NRP report and about the articles cited in it made possible the creation of this teacher's guide. My deep appreciation to them all:

South Carolina State Department of Education Staff: Suzette Lee, Pat Branham, Pam Huxford, Cathy Jones, and Pam Wills

South Carolina Reading Initiative Teaching Team: Diane DeFord, Amy Donnelly, Janet Files, Susi Long, Heidi Mills, and Denise Morgan

South Carolina Reading Initiative Regional Coaches, Professional Development Providers, and/or Interventionist Trainers: Clavis Anderson, Jennifer Brown, Stacie Carlton, Cathy Chapman, Christie Clonts, Robin Cox, Marie Crawford, Diane Dunham, Becca Driggers, Tammy Gacek, Patti Hunnicutt, Lou Jacobs, Jane Johns, Carol Leopold, Sally Nottingham, Pat Smith, Rebecca Thomas, and Jennifer Young

My thanks also to University of South Carolina doctoral student Rebecca Cibic and M.Ed. student Nichole Johnson, who helped check references, M.Ed. student Abbey Mattos, who helped check references and assisted with revisions suggested by reviewers, and anonymous reviewers whose helpful comments informed the final version.

Without all of them, this guide would not have been possible.

Executive Summary

Using Quotes from the NRP Report and This Guide

The National Reading Panel (National Institute of Child Health, 2000) examined the experimental and quasi-experimental research in five areas they considered to be related to reading. This text examines the studies they included and their findings. Their conclusions are outlined below, predominately using quotes from their full report.

Phonemic Awareness

> It is important to note that acquiring phonemic awareness [PA] is a means rather than an end. PA is not acquired for its own sake but rather for its value in helping children understand and use the alphabetic system to read and write. This is why including letters in the process of teaching children to manipulate phonemes is important. PA training with letters helps learners determine how phonemes match up to graphemes within words and thus facilitates transfer to reading and spelling. (P. 2-33)
>
> Teaching children to segment phonemes in words and represent them with letters is the equivalent of invented spelling instruction. (P. 2-41)
>
> Teaching PA with letters "helps children on tests of phonemic awareness and, generally, on tests of word reading." PA does not improve, nor is it intended to improve, reading comprehension (Chapter 2 of this guide).
>
> Teaching phonemic awareness does not ensure that children will learn to read and write. Many other competencies must be taught for this to happen. (P. 2-43)

Phonics

> Teaching students to use phonics skills to read and spell words at the kindergarten level may yield only limited success. (P. 2-126)
>
> There were insufficient data to draw any conclusions about the effects of phonics instruction with normally developing readers above 1st grade. (P. 2-116)
>
> 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 texts and to spell words. (P. 2-116)
>
> Phonics instruction helped regularly achieving first graders be better able to read real and nonsense words. No in-school approach improved the reading comprehension of regularly achieving first graders. (Chapter 3 of this guide)
>
> Three approaches helped at-risk first graders with comprehension: a modified whole language approach . . . , one-on-one Reading Recovery–like instruction . . . , and direct instruction. (Chapter 3 of this guide)

Fluency

Fluency involves being able to “group words appropriately into meaningful grammatical units for interpretation” (p. 3-6, attributed to Schreiber, 1980, 1987). As such, fluency involves reading with “both expression and understanding” (p. 3-11).

Fluency requires the rapid use of punctuation and the determination of where to place emphasis or where to pause to make sense of a text. (P. 3-6)

Vocabulary

Despite the relatively small body of data available, the collective body of research clearly indicates that vocabulary increases with instruction of many different sorts. (P. 4-24)

Estimates of vocabulary size seem to suggest that there would never be sufficient classroom time to instruct students to the level of their acquired vocabulary. This implies that much of a student’s vocabulary will have to be learned in the course of doing things other than explicit vocabulary learning**.** Students may well pick up vocabulary in contexts different from the formal learning of a classroom reading group. It may even be that the vocabulary acquired in this way is more memorable, given the role of motivation in its acquisition because the vocabulary acquired in this way may be far more useful. Repetition, richness of context, and motivation may add to the efficacy of incidental learning. (P. 4-26).

Nagy, Anderson, and Herman (1987) concluded that average reading in grades 3 to 8 probably accounts for one-third of a child’s annual vocabulary growth. (Chapter 5 of this guide)

Comprehension

Reading is an active process of “intentional thinking during which meaning is constructed through interactions between text and reader” (p. 4-39, referenced to Durkin, 1993).

The bulk of the comprehension research has been on strategies which “readers normally acquire . . . informally” (p. 4-5).

There is no information in the NRP report about reading strategies that can be used to improve comprehension instruction for children in kindergarten and first grade. There is very limited information about reading strategy instruction for second-grade children. Teachers need to rely on other sources of SBRR [scientifically based reading research] to help young children better comprehend text. (Chapter 6 of this guide)

The general finding [for grades 3 to 6] is that when readers are given cognitive strategy instruction, they make significant gains on reading comprehension over students trained with conventional instructional procedures. (P. 4-40)

[E]ffective reading instruction is associated more with independent teacher action than with implementation of basal test prescriptions. (P. 4-48)

Strategic reading requires strategic teaching, which involves putting teachers in positions where their minds are the most valued educational resource. (P. 4-49)

1: The National Reading Panel

All across the country, teachers are being told to use practices and materials that are consistent with scientifically based reading research (SBRR). The most often cited source of SBRR is the report written by the National Reading Panel (NRP), a panel commissioned by Congress and convened by the director of the National Institute of Child Health and Human Development in consultation with the U.S. secretary of education to "assess the status of research-based knowledge . . . of various approaches to teaching children to read" (National Institute of Child Health, 2000, p. 1-1).

At the federal level, the mandates come from congressional interpretations of the findings of the NRP report woven into the federal Elementary and Secondary Education Act of 2002, better known as No Child Left Behind, and into the guidance for Reading First grants under No Child Left Behind. At the state and local levels, the mandates come from legislators and administrators who make decisions based on what they believe is in the NRP report.

One problem with the interpretations of the NRP findings is that most policymakers have not read the 449-page report, let alone the 441 studies the report cites.

A second problem is that there are two versions of the report: a full report and a 33-page report, and there are considerable differences between what is reported in the full report and what is reported in the shorter report. Both versions have the same logo on the cover. Both versions are listed as prepared by the NRP and are entitled *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction.* Both versions are available at http://www.nationalreadingpanel.org. Only the 449-page report, hereafter referred to as the full report, has an additional subtitle, *Reports of the Subgroups.* The 33-page report has all the appearances of being the full report. The only place it is identified as a summary is on the link to it on the website. In its paper and downloaded versions, it is not identified as a summary. We know the authors of the full report because they are listed on page iii of the document. We do not, however, know the authors of the shorter report. They are not identified in the document. We only know that the shorter report was not written by the NRP and that it was widely distributed by Widmeyer-Baker, a firm that represents McGraw-Hill, the publisher of Open Court, a commercial reading program (Garan, 2002).

The confusion over "the" NRP report would not necessarily be problematic (just confusing) except that the 33-page report misrepresents the full report on several important points. For example, while the full report says that "phonics instruction taught early proved much more effective than phonics instruction introduced after first grade" (p. 2-93), the shorter report says "phonics instruction produced significant benefits for students in kindergarten through 6th grade" (p. 9). Similarly, while the full report says that "phonics instruction failed to exert a significant impact on the reading performance of low-achieving readers in 2nd through 6th grades" (p. 2-94), the shorter report says, "Systematic synthetic phonics instruction . . . benefits both students with learning disabilities and low-achieving students" (p. 9).

A third problem with the interpretations of the NRP findings is that there is also a video of the NRP findings, also widely distributed by Widmeyer-Baker (Garan, 2002).

Some of the statements in the video also misrepresent what was in the full report. For example, the video reports, “The panel found that systematic phonics instruction is effective for students in kindergarten through sixth grade,” in direct contradiction to the full report. Additionally, the video shows students using Open Court. This, together with the guidance for Reading First grants, has lead some states to erroneously conclude that the NRP found Open Court superior to other reading programs.

These problems and misuses of the research studied by the NRP have eroded the power of teachers to make informed, scientifically based decisions and, as a consequence, have interfered with, rather than supported, effective reading instruction. This is not only problematic but ironic because the full report makes it clear that teachers make the difference, not materials. According to the full report, “Effective reading instruction is associated more with independent teacher action than with implementation of basal text prescriptions” (p. 4-448).

Many people have been critical of the NRP report. See, for example, Allington (2002), Coles (2003), Garan (2002), Krashen (2004), and Yatvin (2000). In the sections that follow, I make reference to the criticisms as appropriate, and while I agree with many of these criticisms, my intent here is not to rehash the criticisms but rather to provide a concise summary of what the report and the research articles did find so that teachers can use this information to counter mandates that are not consistent with the SBRR as reported by the NRP. I do so because, in spite of the limitations of the report, the NRP report—or, more precisely, the perceptions of the NRP’s findings—has informed and is continuing to inform policy. To be heard in the policy conversations, teachers need to arm themselves with data from the report and from the studies cited in the report.

Who Wrote the NRP Report?

Congress specified that the panel should consist of “leading scientists in reading research, representatives of colleges of education, reading teachers, educational administrators and parents” (p. 1-1). The panel ended up consisting of 14 individuals with diverse backgrounds. Garan (2002) reports that panel members “included a certified public accountant, a physics professor, a neuroscientist, an assistant director of the National Science Foundation and seven cognitive psychologists/scientists” (p. 3). Many of the cognitive psychologists/scientists were reading researchers. The panel also included two teachers: Ms. Gwenette Ferguson, who was a middle school reading teacher, and Dr. Joanne Yatvin, a teacher and principal. Dr. Yatvin did not agree with the findings of the panel and wrote a minority report (2000, 1-6). Many people have been critical of the composition of the panel, arguing that more teachers should have been included and that the “reading researchers” and the “representatives of colleges of education” should have included professors representing diverse philosophies about the reading process (such as linguistic, psycholinguistic, and behaviorist perspectives) rather than predominately representing a cognitive psychology perspective.

What Did This Panel Do?

The panel was asked by Congress to “assess the status of research based knowledge including the effectiveness of various approaches to teaching children to read” (p. 1-1). The panel members decided to look into five aspects of reading that they considered

important: phonemic awareness (PA), phonics, fluency, vocabulary, and comprehension. Many people have criticized the panel for their decision to consider only these five aspects of reading. Richard Allington (2005a), for example, who recently served as president of the International Reading Association (IRA), argues that the panel missed five other pillars of reading instruction: classroom organization, matching pupils and texts, access to interesting texts, choice and collaboration, writing and reading, and expert tutoring. In Appendix A of the Minority Report, JoAnne Yatvin (2000) includes two lists of omitted topics—one based on a survey conducted by IRA and the other based on her experiences. The IRA list includes such topics as balanced reading instruction, decodable text, direct instruction, and early intervention. Dr. Yatvin's list includes embedded skills instruction, reading aloud to children, use of predictable texts, and access to literature.

The panel decided to use what they referred to as "evidence-based methodological standards" (p. 1-5). They describe these standards as

> essentially those normally used in research studies of the efficacy of interventions in psychological and medical research. These include behaviorally based interventions, medications or medical procedures proposed for use in the fostering of robust health and psychological development and the prevention or treatment of disease. (P. 1-5)

Many people have criticized the panel for their decision to privilege standards used in "psychological and medical research" (p. 1-5) rather than standards used in educational research.

The panel then divided itself into three subgroups to look at the five aspects of reading they had decided to investigate. Alphabetics looked at the research on PA and phonics. Fluency looked at fluency. Comprehension looked at vocabulary and comprehension. The report consists of findings from 52 PA studies, 38 phonics studies, 50 (fluency) studies on repeated oral reading, 51 (fluency) studies on independent silent reading, 50 vocabulary studies, and 205 comprehension studies. In addition to synthesizing the research on these five elements, the panel also looked at the research on teacher education (32 studies) and computer technology and reading education (21 studies). Because there were so few studies in these last two areas, the panel did not draw conclusions but recommended that additional research be conducted.

It is important to pay attention to the actual number of studies included in the NRP report because the press (and some press releases) reported that the NRP looked at 100,000 studies. The NRP did, indeed, consider a large number of studies for inclusion in their report. However, the number of studies included in the entire report is relatively small (441), and when grade levels and components are considered, the numbers get even smaller. There are, for example, only 5 PA studies involving preschoolers and only 7 phonics studies with kindergartners, 6 of which were conducted with at-risk children. Policymakers need to understand the research base of findings when making decisions about local, state, and federal mandates. Findings from 5 or 6 or 7 studies is a far cry from findings from 100,000 studies.

The panel determined that all subgroups should conduct a meta-analysis of the research in all five areas, and many groups have been critical of this decision to

privilege meta-analysis. However, only the Alphabetics subgroup (PA and phonics) exclusively used meta-analysis techniques. Fluency used it for just 16 studies, arguing that the diverse characteristics of the other studies did not lend themselves to meta-analysis. The comprehension subgroup did not use meta-analysis at all. They did not conduct a meta-analysis on the vocabulary research because they did not consider 50 studies to be "numerically large" (p. 4-17) enough. They believed that meta-analysis was not appropriate for the comprehension studies because "even the studies identified in the same instructional category used widely varying sets of methodologies and implementation" (p. 4-5). These decisions stand, as may be obvious, in stark contrast to the decision of the Alphabetics subgroup to conduct a meta-analysis on 52 PA studies and on 38 phonics studies and the decision of the Fluency group to conduct a meta-analysis on 16 studies.

What Is Meta-Analysis?

Meta-analysis is a statistical technique for looking across studies. To use meta-analysis, there need to be pretest and post-test means, pretest and post-test standard deviations, and a control group. Within these studies, students should be randomly assigned to these control groups. When this happens, the study is considered to be an *experimental* study. When that does not happen, but there still is a control group, the study is considered to be *quasi-experimental*.

A meta-analysis, done correctly, identifies interventions that are statistically significant. In a meta-analysis, a gain effect of .8 (nearly one standard deviation) is considered large, .5 is considered moderate, and .2 is considered small. A significant gain effect indicates that the treatment groups did better than the control groups on assessment measures and that the scores of the treatment group are greater than would have been predicted by chance. In other words, the odds are that it was the treatment, not chance, that accounted for the differences.

Because a meta-analysis involves making comparisons across experimental or quasi-experimental studies, it is important that the students in all the control groups in all the studies use the same material—for example, a basal reader. However, in the meta-analysis that formed the basis for the phonemic awareness and the phonics subreports, the control groups of the studies often differed from one another. Sometimes, for example, a basal was the treatment, and sometimes it was the control. Similarly, sometimes whole language was the treatment, and sometimes it was the control. This concern is sometimes referred to as a problem of *apples and oranges*—one treatment is being compared to apples (e.g., a basal text), and another treatment is being compared to oranges (e.g., whole language). This makes it difficult to draw conclusions about the effectiveness of the treatment as it raises the question of "Effective as compared to what?"

Numerous critiques have been written about the choice of meta-analysis and about how the meta-analysis was done. (See Camilli, Vargas, & Yurecko, 2003, for a re-analysis of the data used in the NRP report. This report was published in a journal edited by Gene Glass, the individual who developed the meta-analysis technique.)

What Did the Subgroups Do?

The subgroups reviewed the literature on the topics they were exploring. Across subgroups, the panel members established criteria for inclusion: The study had "to focus directly on children's reading development from preschool through grade 12" (p. 1-5). It also had to have been published in English in a refereed journal (p. 1-5), a journal whose articles are reviewed by a committee of scholars in the field. The studies also had to be experimental or quasi-experimental studies (p. 1-7) and of "moderate to high quality" (p. 1-7). The subgroups then summarized what they learned from looking across studies.

It is important to understand that while the subgroups adhered to the guidelines established across all studies, each subgroup made additional, and not similar, decisions about which studies to include in their report. These differences involved the characteristics of the children studied. The Alphabetics group, for example, included studies involving children who were and were not "normally achieving" (that is, they included studies with children who were considered to be at risk, learning disabled, reading disabled, etc.). They also included studies conducted with non-English-speaking children; for example, they included a study (Reitsma & Wesseling, 1998) about Dutch-speaking children in Amsterdam being taught by a computer program to blend sounds. The Fluency group included studies on children who were and were not "normally achieving" but, in contrast to the decision made by the Alphabetics group, did not include a study unless it involved children for whom English was their first language. The decisions made by the Comprehension subgroup did not match either of the above decisions: they excluded any at-risk children or any children for whom English was not their first language.

What Did the Studies Find?

In Chapters 2–6, I review the studies cited in the NRP report and draw conclusions about the studies in the Alphabetics, Fluency, and Comprehension subgroup reports. In Chapter 8, I provide an overview of my findings. Tables 1–5, interspersed in the text, and the list of findings in Chapter 5 summarize what the studies cited in the NRP report found. These tables and the list can also be found in the appendix.

2: Alphabetics—Part 1: Phonemic Awareness

Subgroup Members: Linnea Ehri, Ph.D., Chair (professor in cognitive psychology); Gloria Correro, Ed.D. (professor in curriculum and instruction); Timothy Shanahan, Ph.D. (professor in urban education, literacy, and reading); Dale Willows, Ph.D. (professor in human development and applied psychology); and Joanne Yatvin, Ph.D. (teacher and principal)

The Alphabetics subgroup (Ehri, Correro, Shanahan, Willows, & Yatvin, 2000) looked at studies that involved instruction in PA (the ability to focus on and manipulate phonemes in spoken words—such as being able to blend *m/an* to get *man*) and at instruction in phonics (letter-sound correspondences). The subgroup used meta-analysis as their sole tool for determining the impact of PA and phonics training. I examine PA in this chapter, phonics in Chapter 3.

In their report, Ehri et al. (2000) noted that there is a correlation between PA and later reading achievement (p. 2-1 and p. 2-9). That is, young children who have high PA are more likely to end up being good readers. While Ehri et al. acknowledged that there may be a third factor that explains both (e.g., the young children with high PA may have been read to a lot or may already have been readers) (p. 2-30), they chose not to look at studies that examined how children *acquired* PA independent of instruction (as such studies would not be experimental) but chose rather to look only at those experimental and quasi-experimental studies in which children explicitly *were taught* PA skills. They found 52 such studies that met the NRP criteria. The report of this subgroup indicates, however, that all children, including those in the control groups, increased their PA across time. This again raises the possibility that PA may be something that is acquired as a part of the reading process (p. 2-33). Indeed, the panel cautions the reader not to assume that PA skills *cause* reading achievement:

> The biggest source of uncertainty is whether there is a hidden variable . . . [which is] the true cause of the difference; thus the conclusions drawn should be regarded as tentative and suggestive rather than the final word. (P. 2-20)

Ehri et al. (2000) included in their report the findings of 52 studies involving explicit PA instruction. The children in those studies were in preschool through grade 6**:** 5 studies involved preschoolers, 27 involved kindergartners, 14 involved first graders, and 8 involved children in grades 2 to 6. (This totals more than 52 because sometimes children from more than one age group were included in a particular study.) Eleven of the 52 studies were conducted in non-English-speaking countries; 21 studies involved children considered to be at risk. Of the 52 studies, then, 22 involved normally achieving, English-speaking children; all but 2 of these were conducted in the United States.

Ehri et al. (2000) divided the assessments used in the PA training studies into three groups of tests: PA, reading, and spelling. PA tests assessed such skills as segmenting, blending, identifying, reversing, and deleting. Reading assessed such skills

as the ability to identify words and to decode both nonsense and real words. Reading assessments also included time to reach a criterion (how long it took the children to learn something), reading comprehension, reading speed, and miscues. All of these assessments were considered as indications of *reading.* This means that if the meta-analysis showed that teaching blending and segmenting resulted in higher scores on tests of blending and segmenting, Ehri et al. reported that teaching blending and segmenting improved *reading*. This is a distinction that many policymakers do not seem to understand. Too often, policymakers believe, for example, that since the NRP report shows that PA improves *reading*, they should mandate the teaching of PA at all grade levels. What they do not understand is that, many times, training in PA does not improve what most people consider reading (that is, reading comprehension) but rather improves scores on tests of PA. Some studies also show that it helps some students on tests of word reading. Teachers need to be prepared to help policymakers understand the distinction between *reading* as discussed by the NRP Alphabetics subgroup and *reading* as comprehension discussed in school settings. Indeed, as the panel points out, PA is *not* expected to help directly with reading comprehension; it is intended to help indirectly by improving word reading (p. 2-20).

General Phonemic Awareness Findings

Ehri et al. (2000) found that children who were taught the skills associated with PA (e.g., blending and segmenting) did better on related tests (e.g., blending and segmenting) than did children who were not taught those skills. This finding was supported by studies conducted with preschool, kindergarten, and first-grade students. They also found the following results:

> Focusing on one or two skills was better than focusing on multiple skills (p. 2-4 and p. 2-20). This was not tested in any one particular study but is a meta-analysis finding across all 52 PA studies.
>
> Teaching both blending and segmenting led to higher test scores (on such tests as blending and segmenting, deletion, and word identification) than did a multiple skills approach (p. 2-21). As above, this was not tested in any one particular study but is a meta-analysis finding across all PA studies.
>
> Small group instruction is more effective than large group or one-on-one instruction (p. 2-42). This is another finding across PA studies, as opposed to being a finding in any particular study. Teachers, however, are urged to consider what was taught whole group versus small group versus one-on-one, as no PA study looked at the differences among teaching the same content in three different instructional settings.
>
> To teach first-grade children to read words, it is best to teach PA with letters (p. 2-21) (see, for example, Bradley & Bryant, 1983, 1985; Cunningham, 1990; Hohn & Ehri, 1983). Note that when PA is taught with letters, such as through the use of invented spelling, the NRP considered it to be phonics instruction (p. 2-34).

To teach preschool, kindergarten, and first-grade children to read words, it is best to teach PA at the same time the children are receiving reading instruction (see, for example, Brennan & Ireson, 1997; Cunningham, 1990; Solity, 1996).
Teaching PA in isolation does not consistently mean that children will be able to transfer those skills to authentic reading or writing tasks (see, for example, Cunningham, 1990; Solity, 1996).

Findings from Studies Conducted with Preschoolers

As noted above, there were only 5 studies conducted with preschoolers; all the studies were conducted in English. The Fox and Routh (1976) study (which used symbols to teach children to blend) found that children who already knew how to segment benefited from instruction in blending, but that other children did not. This finding is not represented in the broad statements above. Solity's (1996) research with 24 preschoolers is consistent with the conclusion from the meta-analysis that teaching PA is best done as part of or alongside reading instruction (gain effects were .52 for PA and 1.18 for reading words). The three other studies (with 234 preschool children) (Byrne & Fielding-Barnsley, 1991, 1993, 1995; Haddock, 1976; Treiman & Baron, 1983) support another of the meta-analysis findings: children who are taught particular PA skills and then tested on those skills do better on the tests than do children who were not taught those skills. This finding seems like a "common sense" finding. However, it serves as a useful reminder to teachers that when children do not know how to do something, it does not mean that there is something wrong with the children but rather that we, as teachers, need to teach them what they do not know.

Findings from Studies Conducted in English with Normally Achieving Kindergartners

Across the 13 studies conducted in English involving normally achieving kindergartners (Ball & Blachman, 1991; Blachman, Ball, Black, & Tangel, 1994; Brennan & Ireson, 1997; Castle, Riach, & Nicholson, 1994; Cunningham, 1990; Davidson & Jenkins, 1994; Ehri & Wilce, 1987; Farmer, Nixon, & White, 1976; Hohn & Ehri, 1983; Murray, 1998; O'Connor, Notari-Syverson, & Vadasy, 1996, 1998; Tangel & Blachman, 1992; Treiman & Baron, 1983)**,** the NRP report of the Alphabetics subgroup (Ehri et al., 2000) found that children who received PA instruction generally did better on PA tasks than did children who did not get such instruction. This supports the broad finding that when tests measure what children in a treatment group were explicitly taught, the scores of those children are higher than the scores of children who were not taught what was subsequently tested. In most of these studies, normally achieving kindergarten children who received PA instruction also did better on word recognition tasks. This, however, was not true in all cases. The NRP report (Appendix F, p. 2-84) shows, for example, that in the O'Connor, Notari-Syverson, and Vadasy (1996) study, the normally achieving children had only a gain effect of .11 on reading. Similarly, in the Treiman and Baron study (1983), the gain effect of the normally achieving children on a reading task was small: .13.

As noted above, it is also always important to take into consideration what "better" means. Ehri and Wilce (1987), for example, found that students in the treatment group did a "better" job on reading words made up of the onsets and rimes they had

been taught. “Better” meant they began to read, but did not consistently read, 5.5 out of 12 words. Similarly, Hohn and Ehri (1983) found that kindergarten children who were taught to segment using letters did “better” on tasks involving sounds used in training but did not do “better” on sounds not used in training: “[the] performance was close to zero . . . indicating that phoneme segmentation skill does not transfer automatically to longer units that have not been practiced” (p. 757).

Findings from Studies Conducted in English with At-Risk Kindergartners

There were 8 studies conducted in English with at-risk kindergarten children (Brady, Fowler, Stone, & Winbury, 1994; Fox & Routh, 1976; Gross & Garnett, 1994; O’Connor & Jenkins, 1995; O’Connor, Jenkins, & Slocum, 1995; O’Connor, Notari-Syverson, & Vadasy, 1996, 1998; Torgesen, Morgan, & Davis, 1992; Warrick, Rubin, & Rowe-Walsh, 1993). Four of the studies involved researcher Rollanda E. O’Connor (O’Conner & Jenkins, 1995; O’Connor, Jenkins, & Slocum, 1995; O’Connor, Notari-Syverson, & Vadasy, 1996; O’Connor, Notari-Syverson, & Vadasy, 1998). These 4 studies focused on blending and segmenting, as did the Torgesen et al. (1992) study. The conclusion across these five studies is consistent with one of the broad findings from the meta-analysis: children in the treatment groups generally do better on tests of PA (and sometimes on tests of naming and letter-word identification) than do the children in the control groups. In some of the studies, the children with lower skills benefited (e.g., O’Connor, Jenkins, & Slocum, 1995); in others (e.g., O’Connor, Notari-Syverson, & Vadasy, 1996), they did not. Consistent with the pattern that children in phonological treatment groups generally do better on phonological tasks, Warrick et al. (1993) found that language-delayed children who received phonological training did as well on phonological tasks as the normally developing group and that students in the training group maintained this advantage a year later.

The work of Brady et al. (1994) also supports this pattern. In this study of at-risk kindergarten children, those who received PA training did better on PA tests. However, the children in both the treatment and the control groups made comparable gains on reading words and spelling at the end of kindergarten and at the end of first grade. This study raises an important issue about the impact of treatment. Most of the PA studies show the scores of children receiving a treatment are higher than the scores of children not receiving a treatment. The finding of Brady et al. suggests, however, that an increase in scores on a test of PA for kindergartners may not translate into long-term differentiated word-reading and spelling achievement. The NRP report addresses this issue as part of a discussion of the Byrne and Fielding-Barnsley (1993, 1995) study. In that study of children from preschool through second grade, at the end of second grade there was only a “marginal difference in reading comprehension favoring the PA-trained students. However, the 2nd graders did not differ in reading real words or in spelling words” (p. 2-37). The report suggests that the reason the long-term training effects were not stronger is that “the formal reading and spelling instruction . . . in school was sufficiently effective to compensate for the advantage provided by preschool training in PA” (p. 2-37). In reflecting on the impact of treatment, teachers should look for evidence of long-term impact and keep in mind this possibility raised by the NRP: that subsequent reading and writing instruction may in essence “wipe out” any advantage of early PA training.

The Fox and Routh (1984) study also raises some interesting questions. In their earlier study (Fox & Routh, 1976, discussed above), the researchers learned that their intervention helped preschool children who already knew how to segment (gain effect of 1.61) but did not help children who did not already know how to segment (gain effect of −.1). To investigate this issue further, in their 1984 study they tested kindergarten children and identified them as high or low segmenters. They put the low segmenters into three groups: no training, segmenting only, and segmenting and blending. The low segmenters in the segmenting and blending group did better on tests of segmenting, on tests of blending, and on a paired-associate word-learning task than did the children in the other treatment group or in the control group. The low segmenters in this treatment group, however, did not do as well by the end of the training program as the high segmenters had done at the beginning of the study. For example, at the end of the intervention the children in the segmenting and blending treatment group could segment a mean of 41.7 words, while the children in the high segmenting group could segment 43 words before the intervention. More dramatically, the children in the segmenting and blending group could blend 5.10 words after the intervention, compared to the 24.7 words that the high-segmenting group could blend before the intervention. This study raises interesting questions about how it is that the high-segmenting children gained their skill in the first place. Also, as with the Ball and Blachman (1991) study, teachers need to balance out time and effort expenditure with gain. In the Fox and Routh (1984) study, for example, children received additional instruction 15 minutes a day, 5 days a week, for 5 weeks. This improved their ability to both blend and segment but did not bring this ability up to the level of the previously capable peers who received no intervention.

A look across these kindergarten studies indicates that teachers need to think about who gains (and who does not gain) from treatment and for how long. Teachers also need to think about what it means to be “better” and about the amount of instruction necessary to achieve a particular gain.

Studies Conducted in English with First Graders

The NRP included 11 studies, conducted in English, of first graders in their report; 5 of these involved normally achieving first graders (Cunningham, 1990; Farmer et al., 1976; McGuinness, McGuinness, & Donohue, 1995; Uhry & Shepherd, 1993; Weiner, 1994); 7 involved at-risk first graders (Barker & Torgesen, 1995; Bradley & Bryant, 1983, 1985; Hatcher, Hulme, & Ellis, 1994; Hurford et al., 1994; Iverson & Tunmer, 1993; Vadasy, Jenkins, Antil, Wayne, & O'Connor, 1997; Weiner, 1994). (This list totals 12 because the Weiner 1994 study involved both normally achieving and at-risk first graders.)

Findings from Studies Conducted in English with Normally Achieving First Graders

As with the studies for preschoolers and kindergartners, the finding for several of the studies with normally achieving first graders is consistent with one of the general patterns reported by Ehri et al. (2000): generally, children who are taught PA skills do better on tests of PA than do children who are not taught those skills (e.g., Cunningham, 1990; Uhry & Shepherd, 1993; Weiner, 1994). McGuinness et al. (1995), however, found that while first graders in both the treatment groups and in the control group made similar gains on a test of PA (the Lindamood Auditory Conceptualization test), the

children in the treatment groups did better on the word identification and word attack subtests of the Woodcock-Johnson test. Their conclusion is consistent with one of the findings of the Ehri et al. (2000) meta-analysis**:** phonological processing is not sufficient for reading success, and phonological processing activities need to be connected to the alphabetic principle.

Cunningham (1990) also demonstrated that children who received PA training coupled with instruction in how to use PA skills with words while reading (referred to as the meta-level group) scored higher on a standardized reading test at the end of first grade than did children who got only PA training or were in the control group. Another finding from Cunningham's report was that there were no statistically significant differences between the treatment groups of kindergartners in their scores on a standardized reading test. The researcher suggested that this difference in results was because the first-grade children had the opportunity to apply what they were learning (because they simultaneously were getting reading instruction), whereas the kindergarten children did not have this opportunity.

The findings of Cunningham (1990) and McGuinness et al. (1995) are consistent with two other findings of Ehri et al. (2000) about PA training: for first graders, PA is best done with letters and is better done as part of reading instruction rather than separate from it.

While the above researchers were interested in the impact (or lack of impact) of PA on reading, Uhry and Shepherd (1993) were interested in the impact of spelling on reading. One conclusion, which resonates with a broad meta-analysis finding, is that children who were taught to spell words with a consonant-vowel-consonant (CVC) pattern were better able to spell words with a CVC pattern than were the children who did not get such instruction. The children who were taught how to segment also did better on tests of segmentation. There was, however, no difference in reading comprehension scores between the two groups. Weiner (1994) also found that children in the treatment group did better on a segmentation test but did not do better on other measures, including a measure of reading ability. What this suggests is that training in PA may increase children's scores on tasks of PA and on tasks involving word reading but may not impact children's reading comprehension.

There are at least two ways to think about this possibility. Krashen (2001) takes a critical stance. He reviewed PA training studies that included a test of reading comprehension and found 6 such studies, 3 of which were done with children learning to read in English. He noted that the effect of PA training on comprehension was quite low and that in 3 of the 6, PA training had no effect at all. Researchers such as Krashen question the value of PA training. The Alphabetics subgroup (Ehri et al., 2000) took a different stance. While they noted that the impact of PA on comprehension was low (.32), they were not surprised by this finding (p. 2-20) because, they argued, PA is intended to improve word reading, and word reading is just one skill they believe is needed for comprehension. As they explained,

> The task of reading, understanding and remembering information in the text involves multiple processes. Not only must students read the words, but also they must do so rapidly and accurately and must construct meaning across the words and sentences. (P. 2-20)

It is important for teachers to consider their position on the small impact of PA training on comprehension tasks. However, no matter what stance they take, teachers would be wise to keep in mind that the field generally agrees that PA activities are not intended to help with comprehension but instead are intended to help with word reading.

Findings from Studies Conducted in English with At-Risk First Graders

The Alphabetics subgroup (Ehri et al., 2000) included 7 studies of at-risk first graders who received PA training. As has been true at the other grade levels and with normally achieving first graders, several studies (Barker & Torgesen, 1995; Bradley & Bryant, 1983, 1985; Hatcher et al., 1994; Hurford et al., 1994; and Weiner, 1994) found that children who were taught particular PA skills did better on PA tasks than did children not taught those skills. The children in the Barker and Torgesen (1995) treatment group also were able to identify more words on the Woodcock-Johnson subtest on word identification.

In the Bradley and Bryant (1983, 1985) study, children who received PA treatment that included letters did better than those children in the other treatment groups or in the control groups. In the Hatcher et al. (1994) study, the children in the reading-plus-phonology group did better on reading tasks (word identification, word reading, nonword reading, and comprehension) than did children in other treatment groups or in the control group. The children in the reading-plus-phonology treatment group maintained these gains 9 months later. These findings support the conclusion from the meta-analysis that teaching PA for at-risk first graders is best done with letters and as part of reading instruction.

Other studies with at-risk first graders reported findings not tied to the broad patterns of the meta-analysis. Iverson and Tunmer (1993), for example, found that children who were in Reading Recovery and who had added training in phonological decoding did not score higher (on text reading level) than did children in Reading Recovery who did not receive the additional training. Vadasy et al. (1997) found that tutoring was better than no tutoring.

An examination of the first-grade PA studies conducted with English-speaking children makes it clear that these studies support the finding from the meta-analysis that teaching PA with letters and as part of, or alongside, reading instruction helps children on tests of PA and, generally, on tests of word reading. These studies do not support, and are not intended to support, teaching PA to improve reading comprehension.

Findings from Studies Conducted in English with Students in Second Grade and Higher

There were 6 studies included by the Alphabetics subgroup (Ehri et al., 2000) and conducted in English that involved PA training and children in second grade and higher. All the students involved in these studies were considered to be at-risk students. These studies were conducted with unique populations; their findings are not represented in the broad statements from the meta-analysis. Consider these four examples:

(1) Kennedy and Backman (1993) taught a multisensory approach to students, ages 11 to 17, who were labeled as having severe learning disabilities. The treatment

group did not do better than the control group on standardized reading and spelling tests.

(2) Lovett, Barron, Forbes, Cuksts, and Steinbach (1994) provided three types of computer-assisted programs to students who had severe neurological impairments. The students who got the letter-to-sound program (t/r/ai/n) did better on tests of word recognition than did students who got the onset/rime or whole word programs.

(3) Williams (1980) taught severely learning disabled children, ages 7 to 12, a program called ABD (analysis, blend, and decode). Six months after the treatment the only difference between the two groups was that the students in the treatment group did better on a test that required them to say the final phoneme.

(4) Wise, Ring, and Olson (1999) created three treatment groups of children ages 7 to 11 who had reading difficulties. Children in a combination group "explicitly learned how articulatory gestures relate to sounds and spelling-sound patterns and explicitly learned how to manipulate sounds in analytic spelling and reading exercises" (p. 281). Children in a second group learned about articulatory gestures but not how to manipulate sounds, and the reverse was true for the third group. On the post-test, children in all three groups did better on all measures (e.g., word recognition, decoding) than did the children in the control group. One year later, the treatment group did better only on a test of phoneme awareness. The researchers note, however, that the gains in standard scores on word recognition were maintained; the students in the treatment group who previously improved in word recognition as a result of training grew as much during the next year as the students with whom they were matched in word recognition at the end of the previous year.

Findings across the PA Studies

Across the PA studies conducted in English, then, support can be found for the broad patterns identified in the meta-analysis conducted by the Alphabetics subgroup (Ehri et al., 2000). The broad patterns, however, do not apply to all grade levels or to both normally achieving and disabled students. Table 1 provides a concise overview of the PA studies conducted in English. What such a chart misses, though, is the question (and answer) about the appropriateness of PA instruction. According to Ehri et al., PA that leads to reading improvement is best taught with letters. Such instruction is called phonics. According to Ehri et al., PA that leads to reading improvement is also best taught as part of, or alongside, reading instruction. These two statements led logically to the statement that PA instruction, without letters and without a connection to reading, is not likely to lead to improved word reading skills. In other words, there is not a role for PA instruction, defined by Ehri et al. as an oral language activity, in the pre-K to first-grade classroom. What these statements seem to suggest, instead, is that children will learn PA as a part of learning phonics in the context of reading instruction. This finding (no need to teach PA by itself), however, is *not* reported in the NRP document, even though it follows logically from a close look at the SBRR included in the NRP report and at statements made in the NRP report by Ehri et al. (2000). What the Alphabetics subgroup concludes, instead, is that PA helps children to read words and pseudo-words

and also helps with reading comprehension. As they note, “These findings show that teaching children to manipulate phonemes in words was highly effective across all the literacy domains and outcomes” (p. 2-3). To avoid feeling caught between these two lines of thought, both of which come from the Alphabetics subgroup, teachers need to think carefully and deeply about the SBRR included in the Ehri et al. (2000) study and then form and test out their hypotheses about the usefulness of PA instruction in pre-K to first-grade classrooms.

Conclusions from PA Studies

What can be learned from the studies of PA included in the report of the Alphabetics subgroup? Table 1 provides a summary of the instructional implications of the PA studies conducted in English and cited in the NRP report. It seems clear that PA should be taught as a part of helping students learn about sound/symbol relationships and, more broadly, as a part of reading instruction.

Table 1. Instructional Implications from PA Training Studies Conducted in English and Included in the NRP Report

Student level	*If you want to help them with . . .*	*You could teach them . . .*
Pre-K through first-grade students	Specific PA skills	Those skills
First graders	Identifying and reading words	PA skills, using letters, alongside or connected to reading or reading instruction—in other words, teach phonics in context

Helpful Quotes about PA Instruction from the NRP Report

As teachers work out for themselves (and address with others) the instructional implications of the PA studies cited by the NRP, the following comments from the report might prove useful:

> The length of time spent teaching [PA] was influential, with treatments lasting from 5 to 18 hours producing larger [gain] effect sizes than longer or shorter treatments (p. 2-4). . . . [However], the NRP findings should not be translated into any prescriptions regarding how long teachers should spend teaching PA. (P. 2-31).
>
> Exactly how PA instruction should be taught by teachers in their classrooms is not clearly specified by the research. A variety of programs was found to be effective. . . . Ultimately, though, teachers need to evaluate the methods they use against measured success in their own students. (P. 2-4)
>
> Teaching phonemic awareness does not ensure that children will learn to read and write. Many competencies must be acquired for this to happen. (P. 2-43)

It is important to recognize that children will acquire some phonemic awareness in the course of learning to read and spell even though they are not taught PA explicitly. . . . In many of the studies reviewed, control groups showed improvement in phonemic awareness from pretests to posttests, very likely because of the reading and writing instruction they received in their regular classroom. (P. 2-33)

It is important to recognize that children will differ in their phonemic awareness and that some will need more instruction than others. . . . The best approach is for teachers to assess students' PA prior to beginning PA instruction. (P. 2-33)

Teaching children to segment phonemes in words and represent them with letters is the equivalent of invented spelling instruction. (P. 2-41)

Teaching children to invent spellings of words is one way to teach PA. Teachers may need to understand the processes children use to invent spellings, how their spellings become more complete and conventional, and how to promote this growth. (P. 2-44)

Whole language instruction that teaches students to invent spellings by detecting phonemes in words and representing them with letters offers a form of PA training. In Reading Recovery, students may acquire phonemic awareness through the spelling instruction they receive. (P. 2-34)

One instructional activity that is maximally effective for teaching PA in a way that builds a bridge to reading and spelling is that of teaching children to invent phonemically more complete spellings of words. Typically, kindergartners who know letter names or sounds can represent the more salient sounds in words such as beginning and ending sounds, for example, writing B to spell *beaver* or R to spell *arm*. Sometimes their spellings are not conventional, for example, writing Y to spell *wife*. However, the important achievement is that they can distinguish sounds in words. Once they can do this, then teachers can help them detect additional sounds in words and learn conventional spellings for those sounds. (P. 2-39)

It is important to note that acquiring phonemic awareness is a means rather than an end. PA is not acquired for its own sake but rather for its value in helping children understand and use the alphabetic system to read and write. This is why including letters in the process of teaching children to manipulate phonemes is important. PA training with letters helps learners determine how phonemes match up to graphemes within words and thus facilitates transfer to reading and spelling. (P. 2-33)

3: Alphabetics—Part 2: Phonics

Subgroup Members: Linnea Ehri, Ph.D., Chair (professor in cognitive psychology); Gloria Correro, Ed.D. (professor in curriculum and instruction); Timothy Shanahan, Ph.D. (professor in urban education, literacy, and reading); Dale Willows, Ph.D. (professor in human development and applied psychology); and Joanne Yatvin, Ph.D. (teacher and principal)

The Alphabetics subgroup (Ehri et al., 2000) explained in the NRP report that reading is a complex process and that understanding sound-symbol relationships is a part of that process. They noted there has been a lot of debate over how best to teach sound-symbol relationships. To answer this question, they reviewed the literature, looking for experimental studies that had true control groups. They ended up identifying 38 experimental studies, all of which were conducted in English. The studies they found, however, did not all have true control groups because sometimes the same thing, such as a basal reader, was the treatment in one study and the control in another. As with the PA studies, this raises again the question of "better" as compared to what.

Fourteen of the 38 phonics studies cited in the NRP report were conducted with normally achieving children. Only 1 of those studies was conducted with kindergartners, 10 were conducted with first graders, 1 with kindergartners through third graders, and 2 with second graders. These numbers are particularly important because policymakers (and newspaper headlines) have suggested increasing phonics instruction for all students ("Phonics works!"). However, the finding in the NRP report is that phonics only helped first graders. Teachers need to encourage policymakers not to make broad mandates based on such a small data set and not to generalize across grade levels.

Across the 38 studies, the treatments included 17 different approaches the subgroup considered to be *synthetic phonics instruction* (children are taught to convert letters to sounds and then blend them to form words). These treatments included Orton-Gillingham, Direct Instruction, Lippincott, Modified Whole Language, Jolly Phonics, and Sing, Spell, Read and Write (SSRW). A second type of treatment focused on teaching children to blend larger sub-units. Those treatments included Lovett's Analogy strategy and Embedded Phonics. There were 2 treatments considered to be combination programs that are described as direct instruction plus analogy or the reverse, and there were 9 miscellaneous treatments such as those referred to as developmental, whole language, traditional basal, analytic, and spelling mastery. Control groups included a Big Book program, the "regular" curriculum, a basal reader, whole word, whole language, and study skills. What this means is that when Ehri et al. (2000) conclude that *phonics* is helpful, they are including a wide range of treatments—from Orton-Gillingham to Lippincott to whole language—and a wide range of control groups—from basals to whole language to study skills.

Assessments used in the phonics studies included decoding real words, reading nonsense words or pseudo-words, word identification, spelling, comprehension, and oral reading of text. Seventy-six percent of the assessments involved reading or spelling single words; 24 percent involved text reading. All the measures except spelling were considered to be measures of reading. What this means is that when a study showed, for example, that students receiving treatment X did better on reading nonsense words,

the Alphabetics subgroups (Ehri et al., 2000) reported this as evidence that the treatment helped the children as *readers*—that is, it improved *reading*. This is an important distinction for teachers to understand as policymakers sometimes mistakenly believe that phonics helps with all aspects of reading.

It is also important to be cautious when drawing conclusions about the instructional implications of these findings because, as noted above, a wide range of treatments were being compared to a wide range of programs. It is therefore not clear, across studies, what was "better" than what. Consider, for example, the 10 studies conducted with normally achieving first graders. In 6 of these studies, the basal was used as a control group and not as a treatment. In 4 of them, however, the control group was not a basal, and in 1 of them, the basal was both treatment and control. Leach and Siddall (1990), for example, asked parents to read at home with their children in four different ways and studied the impact on reading accuracy and comprehension. Evans and Carr (1985) compared the achievement of a group of students whose teacher used a basal reader to students whose teacher did not. Foorman, Francis, Novy, and Liberman (1991) compared one basal with added phonics to a different basal without added phonics. Klesius, Griffith, and Zielonka (1991) provided one day of professional development to two groups of teachers. One group learned about whole language, and the other group learned about direct instruction. When there is such diversity of control groups, it is not practically possible or statistically logical to compare the findings across the studies and draw conclusions. What would be said, for example, across these 4 studies? What would we say was better than what?

Findings for Kindergartners

There was one study (Bond, Ross, Smith, & Nunnery, 1995–1996) conducted with normally achieving kindergarten students and 6 studies conducted with at-risk kindergarten students (Blachman, Tangel, Ball, Black, & McGraw, 1999; Gersten, Darch, & Gleason, 1988; Martinussen, Kirby, & Das, 1998; Stuart, 1999; Torgesen et al., 1999; and Vandervelden & Siegel, 1997). One of these studies, Stuart (1999), was conducted with kindergartners for whom English was not their first language. As this study is one of only three ELL studies included in the NRP report, broad conclusions about teaching ELL students cannot be drawn from the NRP report. That is, the NRP report does not provide SBRR concerning teaching ELL students.

Based on their meta-analysis, Ehri et al. (2000) concluded that "teaching students to use phonics skills to read and spell words at the kindergarten level may yield only limited success" (p. 2-126). This does not mean that teachers should change what they are doing with kindergartners. What it means is that most of the experimenter-designed treatments were not any better than what teachers were already doing. The findings of 2 of the studies, however, do have instructional implications for phonics in kindergarten. The first comes from Vandervelden and Siegel (1997), who found that inner-city at-risk kindergarten children did better on tests of print matching, spelling, pseudo-word reading, and real word reading when phonics instruction was connected to reading rather than taught as an isolated skill. The second comes from Gersten et al. (1988), who found that children who had an academic kindergarten did better than those who did not.

To understand the treatments in studies whose findings do not inform instruction, consider these 3 studies: Bond et al. (1995–1996), Martinussen et al. (1998), and Torgesen et al. (1999). In the Bond et al. study, regularly achieving kindergarten, first-grade, and second-grade students participated in a phonics program called SSRW. These students were from schools designed as “low stratum” (low socioeconomic status, or SES), “middle stratum” (middle SES), and “high stratum” (high SES). The teachers of the children in the control groups used a basal reader instead of SSRW. On the Woodcock-Johnson letter-identification post-test, kindergarten students from middle-SES schools who participated in SSRW performed better than did their peers in the middle-stratum control group. However, kindergarten children in the treatment group whose school was considered to be high or low SES did not perform better than their peers. On the word attack measure, low-SES kindergartners in the treatment group did better than the low-SES control group, but high-SES kindergartners in the treatment group did not do better than their comparison group. There were no differences on the three other post-tests: Durrell Oral Reading, spelling, and writing. SSRW, then, could not clearly be said to “work” for the children in the treatment group.

In the second study, Martinussen et al. (1998) placed 41 students identified as “demonstrating weakness in successive and phonological processing” (p. 19) into three groups: a successive-phonological group, a meaning-oriented group, and a control group that remained in the regular classroom. Three of the 13 children in the successive-phonological group made gains on a test of phoneme elision; the others did not. Eight of the children in the other groups also showed gains on that test. The successive-phonological treatment, then, was not shown to be superior to the other treatment or to the work by the control group.

In the third study, Torgesen et al. (1999), 138 kindergartners considered to be severely disabled in reading were placed into three treatment groups: phonological awareness plus synthetic phonics, embedded phonics, and a regular classroom group that received tutoring in the skills addressed in the classroom. At the end of second grade, children in the first group (phonological awareness plus synthetic phonics) did significantly better than children in the other groups on tests of phonetic word attack and word identification, but they did not score higher on “real world reading ability”—reading comprehension passages. The authors concluded:

> The ability to construct the meaning of written text is the most important outcome of reading instruction, and we found no evidence that children in the three instructional groups were reliably different from one another on this variable. Without evidence for differences in comprehension, it is not possible to assert that any one of the instructional approaches in this study was ultimately more effective than the other. (P. 590)

Findings for First Graders

There were 11 studies conducted with normally achieving first graders (Bond et al., 1995–1996; Evans & Carr, 1985; Foorman et al., 1991; Freppon, 1991; Fulwiler & Groff, 1980; Griffith, Klesius, & Kromey, 1992; Haskell, Foorman, & Swank, 1992; Klesius et al., 1991; Leach & Siddall, 1990; Leinhardt & Engel, 1981; Snider, 1990) and 10 conducted with at-risk first graders (Brown & Felton, 1990; Eldredge, 1991; Foorman,

Francis, Fletcher, Schatschneider, & Mehta, 1998; Gersten et al., 1988; Mantzicopoulos, Morrison, Stone, & Setrakian, 1992; Marston, Deno, Dongil, Diment, & Rogers, 1995; Santa & Høien, 1999; Traweek & Berninger, 1997; Tunmer & Hoover, 1993; Umbach, Darch, & Halpin, 1989). The Alphabetics subgroup concluded that phonics instruction helped regularly achieving first graders be better able to read real and nonsense words. No in-school approach improved the reading comprehension of regularly achieving first graders although asking parents of first graders to do paired reading or provide direct instruction improved reading comprehension more than asking parents to just listen or to pause, prompt, and praise (Leach & Siddall, 1990). Three approaches helped at-risk first graders with comprehension: a modified whole language approach (Eldredge, 1991), one-on-one Reading Recovery–like instruction (Santa & Høien 1999), and direct instruction (Umbach et al., 1989).

Findings for Normally Achieving First Graders

Some of the studies conducted with normally achieving first graders involved treatments that were shown not to lead to differentiated performance for all the children who received the intervention (e.g., Bond et al., 1995–1996; Evans & Carr, 1985; and Klesius et al., 1991). As detailed in the previous section about kindergartners, Bond et al. (1995–1996) provided a program called SSRW for regularly achieving kindergartners, first graders, and second graders. Low-stratum first graders in the treatment group did better on tests on work attack than did their peers in the control group. The first graders in the other two strata did not. This pattern, of the treatment working for some groups and not others, also characterized the letter-identification task. Evans and Carr (1985) compared students who received a basal program (that is, the basal was the treatment) with children in a British Infant School in which no basal was used. The researchers did not pretest the children but assumed they would be similar because they were all from the same socioeconomic neighborhood. At the end of a year, the children in the basal classroom did better on cloze tests and answering questions about passages in the primer. There was, however, great variability, and in some classrooms the basal children did worse than the children in the British Infant School. Both groups did equally well answering questions about grade 1 and grade 2 passages. The Klesius et al. (1991) study provides a third example. One group of teachers was given a half-day inservice on direct instruction, and another group was given a day-long inservice on whole language. At the end of the year, the researchers compared the scores of their students on several measures and found no differences between the two groups.

Another study “worked” for word attack and spelling, but not for comprehension. In that study, Snider (1990) provided direct instruction to two classrooms of first graders, most of whom were learning English as a second language, and compared their scores to the Iowa Test of Basic Skills (ITBS)’s scores for first grade the previous year. The direct instruction group, as noted above, did better on word attack and spelling than did first graders the year before, but they did not do better than those children on reading comprehension.

The remaining studies involved treatments that were shown to be better than the control group across most measures. Teachers should exercise caution when interpreting these findings, though, as the control groups varied considerably. Teachers

should therefore consider carefully the reported success of the program relative to the comparison group.

Foorman et al. (1991) compared the growth of children in a public school using Harcourt-Brace with children in a parochial school using Scott-Foresman and additional phonics readers and worksheets. The NRP considered the Scott-Foresman group as the treatment group. The researchers found that the children in the treatment group more rapidly increased their ability to spell and read words correctly. Because of the design, however, it is not possible to determine if the difference between the two groups was due to the different basal used by each group or to the additional phonics provided to one group.

Freppon (1991) selected 24 children who scored in the sixth, seventh, and eighth stanine, 12 from literature-based/whole language classrooms and 12 from skills-based/traditional classrooms. Children (1) were interviewed, (2) were given altered passages (from Canney & Winograd, 1979) and asked if this was something that could be read (why? why not?), (3) were asked to read it if they said it was something that could be read, and (4) were asked to read a story. The altered passages consisted of one intact text, one passage of letter strings, and three passages of incomprehensible text (To bus was it Mrs. the go the) (p. 148). On the basis of the interview, the researcher reported:

> The skill-based group and the literature-based group were similar in that all the children expressed an interest in knowing and learning words and exhibited an understanding of the use of phonics or sounding out words as a reading strategy. However, the two groups differed in that the literature group expressed understanding of (a) varied reading strategies, (b) metacognitive knowledge and (c) a view of reading as a language process which involves meaning-making. (P. 155)

On the altered passages, both groups accepted the intact passage and rejected the letter string package. The children in the literature group rejected the passages more often and for different reasons. For example, 92 percent of the children in the literature group and 66 percent of the children in the traditional group rejected the passages because they did not think the words were “true” or “real”; however, 33 percent of the children in the traditional group rejected the passages because “the words were hard or because they didn’t know the words” (p. 156).

The errors that the children made while reading were analyzed. The children in the literature-based group more often used phonics in concert with meaning and grammar to figure out new words (34 percent of the time) than did the children in the traditional program (8 percent). The children in the traditional program used only phonics more often (60 percent) than did the children in the literature-based group (24 percent). The children in the literature group were more often successful (53 percent of the time) in their attempts to decode unfamiliar words than were the children in the traditional group (32 percent). The researcher concluded:

> Without sound understanding of reading as a meaning-making process, that is, without having the big picture, children cannot benefit from being taught about the abstract or graphophonemic inner workings of print. (P. 160)

Fulwiler and Groff (1980) compared the end-of-first-grade test scores of 73 children whose classroom teachers had used an intensive phonics program (Lippincott) with the test scores of 74 children whose classroom teachers had used a sight word program (Cop-Clark Canadian Reading Development) series. All the children were "judged" to have "the same level of intelligence . . . based on observations of the socioeconomic backgrounds of the pupils involved and on the intelligence test scores of the other children in the schools the subjects of the present day [*sic*] attended" (p. 52). The authors found that the children whose teachers used Lippincott had statistically significantly higher scores on tests of vocabulary, word analysis, and reading comprehension.

Griffith et al. (1992) compared first graders in whole language and traditional (basal) classrooms on measures of spelling, writing, decoding, and comprehension. There were no significant differences between the two groups. In this study, a test was given in the fall to determine which children had high PA and which children had low PA. The researchers found that at the end of the year, the high PA children outscored the low PA children, regardless of whether the children were in the treatment or control group. The three exceptions were that the low PA children in the whole language classroom were able to spell more unpredictable words than were their counterparts in the traditional classroom. The high and low PA children in the whole language classroom increased the number of words they used in writing while the children in the traditional classroom did not. The low PA children in the traditional classroom were more accurate at spelling parts of words.

Leach and Siddall (1990) asked four groups of parents to engage with their children at home in four different ways: hearing reading (listening to children read); paired reading; pause, praise, and prompt; and direct instruction. All parents spent the same amount of time with their children. For both reading accuracy and reading comprehension, there were no significant differences between the paired reading and direct instruction groups, but there were differences between these two groups and the other two groups.

Leinhardt and Engel (1981) conducted an evaluation of the implementation of the New Primary Grades Reading System (NRS) and concluded that students in the treatment groups outperformed students in the control groups (who used a variety of basal readers) on a variety of tasks—letter names, Stanford Achievement, Murphy-Durrell Reading Readiness analysis.

As must be clear, the findings are diverse, as are the treatments and the control group. The only cross-study statistical comparison that can be made involves the three studies in which treatments for normally achieving first graders were compared to instruction using a basal reader.

Table 2. Instructional Implications from Phonics Studies Cited in the NRP Report When Treatment Is Compared to a Basal

If you want to help normally achieving first graders with . . .	*You could . . .*
Better knowing letter names, having higher scores on reading readiness and on a standardized reading test than did children using a basal	Use the New Primary Grades Reading System (Leinhardt & Engel, 1981)
Understanding that reading is a meaning-making process, using phonics more effectively to figure out unknown words, and having greater metacognitive understanding of reading than did children using a basal	Put them in a literature-based/whole language classroom rather than in a traditional/basal-based classroom (Freppon, 1991)
Spelling more unpredictable words and increasing the number of words they use in writing than did children using a basal	Put them in a whole language classroom rather than a traditional classroom (Griffith et al., 1992)

In other words, for normally achieving first graders, three interventions (a whole language classroom, a literature-based classroom, and the NRS) were shown to be "better" than a basal. This is very important for teachers to understand, particularly teachers in districts that insist on the adoption of a core reading series. There is nothing in the NRP report that suggests that a core reading series provided more help for first graders than did approaches that did not involve a basal reader. This finding resonates with a finding by Harste (1985). In his meta-analysis of reading comprehension research in the years 1974–84, the researchers in his study found that nearly everything other than the basal had a gain effect score of about .7.

This table also makes clear that one of the three interventions that was "better" than the control was an intervention that had an emphasis on phonics, and the other two were consistent with a whole language approach in which phonics, meaning, and grammar are addressed simultaneously. The NRP report, then, did not settle what is sometimes known as the "reading wars." What the studies cited by the NRP show instead is that different approaches (phonics emphasis versus whole language) accomplish different goals.

Findings for At-Risk First Graders

As with the phonics studies with kindergartners and with normally achieving first graders, some of the interventions with at-risk first graders were not successful or had success in some areas but not in others. Brown and Felton (1990), for example, compared two different types of basal instruction (Lippincott and Houghton-Mifflin) in first and second grade for children considered in kindergarten to be at risk of reading failure. At the end of second grade, there were no statistically significant differences between the two groups on any of the measures. Foorman et al. (1998) arranged for three groups of first and second graders to participate in one of three phonics-based

programs: (1) direct instruction of letter-sound correspondences practiced in decodable texts—an Open Court basal reader, (2) less direct—instruction in onsets and rimes embedded in connected text, and (3) implicit—alphabetic code taught as part of reading text. The implicit group was considered to be the control group and was characterized by the researcher as a whole language group. The children were all receiving Title I services and were involved in the intervention for 30 minutes a day for a year. At the end of first grade, the children in the direct instruction group did better on tests of word recognition (gain effect of 1.63) and word decoding (1.14) than did the control group. The less direct group also did better than did the control group (.56 and .51, respectively). Both groups also did better on spelling (gain effects of .56 and .26) and on comprehension, although the gain effect scores for comprehension were low (.32 for direct instruction and .1 for indirect). A year later, the direct instruction group still did better on word identification and decoding. However, there was a negative gain effect on spelling and on comprehension. That is, the children in the direct instruction group did worse on a comprehension measure than did children in the less direct or implicit groups. For comprehension, the gain effect for direct instruction compared to the control group was −.19, and the gain effect for less direct instruction compared to the control group was −.24. In other words, at the end of two years, the children in the control (whole language) group did better on comprehension measures than did the children in the treatment (Open Court) group.

In a study of a third intervention that did not “work,” Mantzicopoulos et al. (1992) compared the progress of three groups of first graders identified, in kindergarten, to be at risk for reading failure. One group participated in a one-on-one tutoring program called TEACH, which focused on perceptual remediation. A second group got tutoring in phonics (10 minutes), read aloud (10 minutes), and spelling (10 minutes). The third group got no additional support and served as the control group. The researchers found that, overall, children in the phonetic tutoring group had higher word attack scores at the end of two years than did the children in the TEACH program or in the control group. However, “children with more severe perceptual difficulties did not profit from either intervention” (p. 583).

There were no differences between the three groups on a standardized reading test. The researchers concluded that “one-to-one tutoring is not likely to result in comprehension achievement gains if its focus is on narrow and isolated instructional activities” (p. 573). They quote Adams (1990):

> [C]omprehensive gains are not likely to be observed when phonics instruction occurs in a vacuum, without consideration for the structures of language, the functions of written text and its personal meaning for the student. (P. 584)

Marston, Deno, Dongil, Diment, & Rogers (1995) tested out “six research-based teaching strategies with 37 special education resource teachers and 176 students with mild disabilities” (p. 20). They concluded that the results were inconsistent. The Traweek and Berninger (1997) study was another study that consistently did not show superior findings for the treatment group. In their study, they compared the impact of DISTAR (a direct instruction program) to the impact of an Integrated Reading-Writing approach. Both groups performed similarly on the word reading subtest of the Stanford

Diagnostic Reading and on writing samples. The only reported difference between the two groups was that the children in the DISTAR program “tended to acquire only subword (orthographic-phonological) connections” while the children in the integrated approach made those connections but at the “whole word and subword level” (p. 160).

Three studies showed an impact of the treatments, but the control groups were so unlike the treatment groups that it is not possible to determine which aspects of the treatment were responsible for the differences in the scores. In the first of these studies, Santa and Høien (1999) provided 30 minutes of tutoring daily to 23 children considered to be in the lowest 20 percent of their class. Tutors who took a graduate course at the university provided instruction that the authors consider to be similar to Reading Recovery with added explicit instruction in phonological analysis. Children in the control group spent 30 minutes reading in small groups. They received no instruction in word study; “[p]ractically the entire 30-minute session was spent on reading” (p. 64). The children who received the tutoring did better on the word recognition, nonword reading, and passage reading subtests of the Woodcock Johnson Reading Mastery Test. In thinking through this study, teachers need to raise the question “Compared to what?” One group of children got one-on-one tutoring; the other group got no instruction and read in a small group setting. It is not clear whether the benefit came from tutoring (versus no tutoring) or instruction (versus no instruction).

In the second study, Umbach et al. (1989) investigated the success of a direct instruction program as compared to a traditional basal program. The direct instruction was provided by master’s degree students participating in a practicum. Those students were provided training in direct instruction and in remediation strategies throughout the school year. University supervisors made weekly observations and taped lessons that were then critiqued. The supervisors also sometimes took over the classes and provided demonstration lessons. The teachers of the basal program had been teaching in the rural district for 10 years and “were given no additional training but were expected to teach the group as they normally would” (p. 3). At the end of a year, students in the direct instruction group scored higher than did control group students on the Word Identification subtest of the Woodcock-Johnson Reading Mastery Test. Because the individuals who provided the tutoring were so different from one another (graduate students versus teachers who had been teaching in a rural district for 10 years), it is not possible independently to determine the effect of the intervention.

In the study by Eldredge (1991), three first-grade teachers in classrooms considered to be whole language classrooms added to their daily schedule 15 minutes of direct instruction in phonics. These children were compared to children in classrooms not considered to be whole language and in which a basal was used. Children in the modified whole language classrooms had greater gains in vocabulary, comprehension, and phonics assessments. It is possible, however, that the difference between the two classrooms represents not a difference between additional phonics and no additional phonics but instead a difference between whole language classrooms and basal classrooms. The design of the study does not allow an interpretation of the effect of the added phonics by itself.

If, however, the muddiness of these studies is not taken into account, the chart for phonics studies of at-risk first graders who received a treatment that was not from a

basal and whose scores were compared to children who received “regular” classroom instruction still would only have three entries, as shown in Table 3.

Table 3. Instructional Implications from Phonics Studies (with At-Risk Students) Cited in the NRP Report in Which the Treatment Is Compared to a Basal

For at-risk first graders	
If you want to help at-risk first graders with . . .	*You could . . .*
Vocabulary, comprehension, and phonics (as compared with children whose teacher uses a basal)	Offer explicit phonics instruction as part of whole language instruction (Eldredge, 1991)
Word identification and passage comprehension on the Woodcock-Johnson (as compared to children whose teachers were from the district, used a basal, and were provided with no additional support)	Have master’s degree students offer direct instruction and provide university-based supervisors who help the students weekly across the year (Umbach et al., 1989)
Word recognition, nonword reading, and passage reading subtests of the Woodcock-Johnson (as compared to children who read books with their teacher in a small group)	Have graduate students provide 30 minutes of one-on-one Reading Recovery–like tutoring (Santa & Høien, 1999)

Findings for Second through Sixth Graders

There were 14 studies (Bond et al.,1995–1996; Foorman et al., 1998; Foorman et al., 1997; Gittelman & Feingold, 1983; Greaney, Tunmer, & Chapman, 1997; Lovett et al., 1989; Lovett & Steinbach, 1997; Lovett, Warren-Chaplin, Ransby, & Borden, 1990; Lum & Morton, 1984; Marston et al., 1995; Silberberg, Iversen, & Goins, 1973; Vickery, Reynolds, & Cochran, 1987; Wilson & Norman, 1998; and Lovett et al., 2000, cited as “in press” in the NRP report) cited in the NRP report that were conducted with second through sixth graders**.** Only 3 of the studies included children who were considered to be normally achieving: Bond et al. (1995–1996), discussed above; Lum and Morton (1984), who studied the impact of a spelling program on spelling; and Vickery et al. (1987), who studied the impact of a phonics program (Orton-Gillingham) on students in a remedial program who received only the treatment versus regular classroom instruction to which the treatment was added. The other studies were conducted with at-risk, reading- or learning-disabled students. Ehri et al. (2000) concluded 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. (P. 2-116)

Across all the studies with these students, Ehri et al. (2000) concluded that “phonics programs did not produce significant growth in reading comprehension” (p. 2-

116). These findings are very important ones for teachers as they often are asked, in second grade and higher, to focus instruction on phonics. They would be wise to point out that the SBRR in the NRP report suggests that such instruction will not help either normally achieving or at-risk students in second through sixth grades as readers.

Looking at Types of Phonics Programs

The Alphabetics subgroup (Ehri et al., 2000) examined the impact of different types of phonics programs (e.g., synthetic versus subunits) and concluded that the research showed no differences between the approaches (p. 2-119). What this means for first grade (the only grade level for which phonics instruction was found to be effective) is that the phonics treatments that "worked" ranged from whole language (Freppon, 1991; Griffith et al., 1992), modified whole language (Eldredge, 1991) and Early Steps (similar to Reading Recovery) (Santa & Høien, 1999) to phonics-based programs such as NRS (Leinhardt & Engel, 1981) to direct instruction (Umbach et al., 1989). Each of these treatments, though, "worked" in different ways. Teachers would be wise to revisit Tables 2 and 3 of this guide in light of their instructional goals and consider as effective those interventions that might help them reach those goals. In dealing with policy matters, what teachers need to help others understand is that there is not one program that is better than any other.

Across the phonics studies, then, the findings show that both whole language and direct instruction approaches helped normally achieving first-grade students do "better" on the assessments used in the studies (Freppon, 1991; Griffith et al., 1992; Leinhardt & Engel, 1981). These interventions did not, however, help these students with comprehension. A direct instruction program (Umbach et al., 1989) helped at-risk first graders improve their word attack skills and their comprehension, as did a Reading Recovery–like one-on-one intervention (Santa & Høien, 1999) and phonics plus whole language (Eldredge, 1991). Phonics instruction was not useful at all for kindergarten students or for students in grades 2 to 6.

Helpful Quotes about Phonics Instruction from the NRP Report

As teachers work out for themselves (and address with others) the instructional implications of the phonics studies cited by the NRP, the following comments from the report might prove useful:

> Programs that focus too much on the teaching of letter-sound relations and not enough on putting them to use are unlikely to be very effective. (P. 2-96)
>
> [S]ystematic phonics instruction should be integrated with other reading instruction to create a balanced reading program. (P. 2-97)
>
> Phonics instruction is never a total reading program. . . . 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. (P. 2-97)

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. (P. 2-117)

By emphasizing all of the processes that contribute to growth in reading, teachers will have the best chance of making every child a reader. (P. 2-97)

Please note: The above comments align with the conclusion of a re-analysis of the NRP data done by Camilli et al. (2003), which reported that the NRP data showed it was best to teach sound/symbol relationships in language-rich settings. This is also consistent with a comment Mantzicoupolos et al. (1992) attribute to Adams (1990): "Comprehensive gains are not likely to be observed when phonics instruction occurs in a vacuum, without consideration for the structures of language, the function of written text and its personal meaning for students" (Mantzicoupolos et al., 1992, p. 584).

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. (P. 2-111).

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. (P. 2-113)

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. (P. 2-97)

Teaching students to use phonics skills to read and spell words at the kindergarten level may yield only limited success. (P. 2-126)

There were insufficient data to draw any conclusions about the effects of phonics instruction with normally developing readers above 1st grade. (P. 2-116)

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. (P. 2-116)

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. . . . 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? . . . [T]he latter is clearly preferable. (P. 2-97)

4: Fluency

Subgroup Members: S. J. Samuels, Ed.D., Chair (professor in educational psychology); Timothy Shanahan, Ph.D., Co-chair (professor in urban education, literacy, and reading); and Sally Shaywitz, M.D. (pediatric physician)

The authors of this report (Samuels, Shanahan, & Shaywitz, 2000) considered fluency to be a "critical component of skilled reading." They noted that "fluent readers can read text with speed, accuracy, and proper expression" and that "fluency depends on well developed word recognition skills." They cautioned, however, that such skills "do not inevitably lead to fluency" (p. 3-1). The recent concern for fluency came about, the authors explain, because of a disproportionate amount of attention that had been given in the past to having students get the words "right" while sacrificing meaning. This emphasis encouraged word-by-word reading, which made comprehension difficult. The renewed interest in fluency is an attempt to marry students' abilities to read words with their abilities to make sense of text.

It is important for teachers to note that this current conceptualization of fluency consequently includes more than just rate. This matters because, in various schools across the country, teachers are being told—reportedly on the basis of the NRP report—to encourage their students to read faster and, indeed, are being asked to time the number of words students can read in one minute. However, according to the NRP fluency authors and to the field more broadly, fluency involves being able to "group words appropriately into meaningful grammatical units for interpretation (Schreiber, 1980, 1987)" (p. 3-6). As such, fluency involves reading with both "appropriate expression and understanding" (p. 3-11). Hopefully, policymakers who understand that fluency includes grouping words appropriate for meaning will no longer require teachers to encourage fast reading while ignoring comprehension. In one elementary school, for example, when students were given a timed passage under high-stakes conditions, they read as many of the words as they could in one minute—skipping all the words they did not know. The paragraph that follows this one might then be read as the following:

> The of this with reading as a was to the of two to the first reading reading reading reading readers out here that their of reading did not reading which has been as of and of time its use has been found to have little or no to in reading the the of or reading.

Teachers need to help policymakers make better choices and consistently include meaning making as a part of fluency instruction and fluency assessment.

The intent of this subgroup concerned with fluent reading as a comprehension tool was to examine the effectiveness of "two major instructional approaches to fluency development." The first involved oral reading (e.g., repeated reading, neurological impress, radio reading, paired reading, readers theater). Samuels et al. (2000) pointed out here that their definition of oral reading did not include "round-robin reading . . . which [has] been criticized as boring, anxiety provoking, disruptive of fluency and wasteful of instructional time. . . . [Its] use has been found to have little or no

relationship to gains in reading achievement" (p. 3-11). The second instructional approach involved increasing the amount of independent or recreational reading.

Samuels et al. (2000) began their work by reviewing the literature on oral reading and concluded there were too many studies. They decided to limit their review to studies conducted since 1990. This gave them 346 studies, which they reduced to 77 studies after eliminating those that did not meet their criteria of being refereed, conducted with English-speaking students, and on the topic of fluency. The authors later used the references from these 77 to add another 21 studies to their collection—for a total of 98 studies. If certain information was omitted (e.g., sample selection criteria), the study was eliminated. Samuels et al. ended up with 50 studies. Only 16, all of which fell into their Group Experiments category, were used in the meta-analysis. The students involved in the fluency studies were in kindergarten to grade 12. The 50 studies were divided into four categories: Immediate Effects (14 studies), Group Experiments (16 studies), Single Subjects (12 studies), and Methods Comparison (8 studies). Across all categories, the assessments included tests of word knowledge, comprehension, and fluency. The meta-analysis of the 16 studies within Group Experiments, coupled with their summarization of the findings from the other 34 studies, led the Fluency subgroup to conclude that repeated oral reading helps with word recognition (.55), fluency (.44), and comprehension (.35).

However, within each category there were limitations of the data that the authors detailed. They noted, for example, that the Immediate Effects studies looked only at immediate effects (the ability to read fluently text that students had just practiced), and so there was no information available about the degree to which these interventions helped students with subsequent texts or with subsequent rereadings of the same text. They concluded that the Group Experiments studies did not provide "sufficient data to allow for a sound analysis of the relative impact of repeated reading procedures on students at different grade levels" (p. 3-17). The students in the Single Subjects studies were all conducted with students considered to be "low level" readers, and all but one were one-on-one interventions. The authors concluded that "either of these factors could magnify the effect" (p. 3-19). None of the studies in the Methods Comparison group had a true control group, and so the authors concluded that it was "not clear whether [the] gains were greater than expected in the amounts of time studied" (p. 3-19).

A few examples will perhaps make these findings more clear. Consider first the Immediate Effects studies. In one of them, Neill (1980) asked "disabled secondary students" to reread passages under timed conditions. Neill reported that one student "whose original time was 175 seconds reduced that to 25 seconds with 100% of the words pronounced correctly in 20 trials. He continued to read the same passage at least twice each day for 17 school days" (p. 64). In a second study, Sindelar, Monda, and O'Shea (1990) asked two groups of students (one group was learning disabled; the other group was not labeled) to reread passages at their instructional reading level or at their mastery level. The fluency and comprehension of both groups improved. In a third Immediate Effects study, Rasinski (1990) asked one group of third graders to engage in repeated readings and another group to participate in a listening from reading treatment. There were no differences in the fluency of the two groups. All of these studies, though, involved rereading the same passage, and thus Samuels et al. (2000)

concluded that no findings could be stated about the impact of these measures on texts not included in the study.

There were 16 studies that Samuels et al. (2000) classified as Group Experiments. A meta-analysis was conducted on these 16 studies. The 605 students participating in these studies were normally achieving students in second through fourth grade and poor readers in grades 2 to 9. There were several different kinds of interventions. Eldredge (1990), for example, looked at the reading gains made by third graders with low reading scores who were given small group instruction versus those who were not. As a part of the assisted instruction, the teacher read parts of the book out loud, and then, at the end of the instructional period, students read the story in dyads without teacher assistance. Students who received the assisted reading instruction had greater reading gains than those who did not receive it. The method used by Hollingsworth (1978) involved a very different method—impress reading—in which the teacher and the student read the text aloud simultaneously. There were no differences in reading achievement between the students in the treatment and in the control group. Lorenz and Vockell (1979) also used the impress method and similarly found no differences between the treatment and control groups. In a fourth study in this category, Rasinski, Padak, Linek, and Sturtevant (1994) asked second graders to engage in reading a passage chorally as part of a fluency development lesson (FDL). The increase in their reading level was compared to students who did not participate in this FDL. The oral reading rates of the students in the treatment group were significantly higher than those of the students in the control group. There were no differences between the prereading and post-reading levels of the two groups of children. A final example: Shany and Biemiller (1995) established two treatment conditions—teacher-assisted and tape-assisted—and found that the 19 third and fourth graders did not differ from each other on gains of accuracy and comprehension.

There were 16 studies included in this category and thus in the meta-analysis. The studies varied not only by method but also by grade level and by type of student involved. Although this subgroup did not provide an appendix with grade level and descriptive information about the students, I was able to find and read 10 of the 16 studies. Of those 16, none involved kindergartners or first graders; 2 involved low-achieving second graders; 2 involved regularly achieving third graders; and 3 involved low-achieving third graders. There were also 2 studies of regularly achieving fourth graders and 3 of low-achieving fourth graders. There were 4 studies of low-achieving fifth graders. (The number does not add up to 12 as some studies, e.g. Simmons, Fuchs, Fuchs, Mathes, & Hodge, 1995, involved students from more than one grade level.)

Because the methods varied so considerably from each other and because there were so few studies at any grade level, the Fluency subgroup (Samuels et al., 2000) concluded that no conclusions could be drawn about the effectiveness of a given treatment with a given age level. They did, however, conduct a meta-analysis of this research, and the broad conclusions drawn by this group are based, in part, on that analysis.

The third group of studies—Single Subjects—involved students whose reading was considered “low-level.” In the study, for example, by Blum et al. (1995), students for whom English was not their first language received tape-recordings of familiar books to

be played at home. The 5 students all increased their reading fluency. Gilbert, Williams, and McLaughlin (1996) asked 3 learning-disabled students to participate in a study that involved having the students listen to a taped passage, read it along with the teacher, and then read independently. They reported that this treatment increased the number of words read correctly and decreased the number of words read incorrectly. In a third study, Kamps, Barbetta, Leonard, and Delquadri (1994) provided peer tutoring to 3 "high functioning students with autism" and found that the tutoring increased fluency and correct responses to comprehension questions (p. 49). Samuels et al. (2000) note that the findings from these studies are a close match to the findings from the meta-analysis but express concern that because these studies were exclusively conducted with low-achieving students and almost all were conducted one-on-one, the effects of the treatment might have been magnified.

There were 8 studies in which methods were compared. However, as noted by this subgroup, none of the 8 included true control groups. Carver and Hoffman (1981), for example, looked at the effect of repeated readings on high school students and concluded that repeated readings might be more helpful for "beginning level readers" and might not be effective with students above fifth grade. Dixon-Krauss (1995) paired first and second graders for partner reading and found that both groups of students improved in word recognition and "accurate evaluations of their own reading progress." There was "minimal improvement in fluency" (p. 45). Dowhower (1987) assigned transitional second-grade readers to two repeated reading treatment groups—read-along or independent practice. Both groups improved their "rate, accuracy, comprehension and prosodic reading" (p. 389). However, because these studies did not include a control group, the authors of this subgroup concluded that no conclusions could be drawn about the effectiveness of the treatments used. Across the oral reading studies, then, there is some evidence that repeated readings help with word recognition and reading rate "with the smallest effects evident in reading comprehension" (Samuels et al., 2000, p. 3-18).

In addition to reviewing this literature, Samuels et al. (2000) also examined the research on studies that examine the effect of encouraging students to read more. The authors found only 14 studies that examined this question using an experimental design. Because the number was so small, Samuels et al. decided not to conduct a meta-analysis of these studies. In the report, instead, the authors detailed the findings from each of these studies. They noted that, for the most part,

> these studies found no gains in reading due to encouraging students to read more. It is unclear whether this was the result of deficiencies in the instructional procedures themselves or to weaknesses and limitations in the study designs. (P. 3-27)

Samuels et al. concluded that "it is not that studies have proven that this [independent reading] cannot work, only that it is yet unproven" (p. 3-27). The authors suggested:

> To really understand the implications . . . it is important to compare these routines against procedures in which students actually read less. . . . Without such information, one might only be comparing the effects of different forms of

> reading practice rather than comparing differences in amount of reading practice. . . . [N]one of these studies could even demonstrate that they clearly increased the amount of student reading because none of them measured an adequate baseline of current or previous reading engagement. (p. 3-27)

In other words, Samuels et al. reported that there was not enough experimental or quasi-experimental research for them to draw a conclusion about independent reading. (There is, however, correlational data on this topic cited in an earlier federal report on reading, *Becoming a Nation of Readers* [Anderson, Hiebert, Scott, & Wilkinson,1985]. That report concludes: “Research suggests that the amount of independent, silent reading children do in school is significantly related to gains in reading achievement” [p. 76].)

It is important for teachers to understand that scarcity of experimental or quasi-experimental research in this area often has to do with ethics. For example, in order to conduct experimental research, researchers would have to randomly assign some students to a group that was encouraged, permitted, or helped to read more extensively than they allowed students in another group to read. Because the field generally believes that that there is a relationship between time spent reading and reading achievement (see, for example, Allington, 1977), restricting access to reading would be considered by many to be unethical. There are similar difficulties in conducting quasi-experimental research, as the research team would have to identify an existing group of students (e.g., particular classrooms) and provide opportunities for them to read more in and out of school and then limit the opportunity for a matched group of students to do the same.

In conducting this review of the literature, Samuels et al. (2000) made a number of decisions about how to proceed that differed from the decisions made by the Alphabetics subgroup. One decision was to narrow their search to particular types of interventions (repeated oral reading and independent reading). Because they narrowed their search in this way, they did not face the problem faced by the Alphabetics subgroup, which had a wide range of interventions. A second decision made by Samuels et al. was not to conduct a meta-analysis on all the studies but to qualitatively read across the studies and draw conclusions based on expert opinion. As a result, Samuels et al. were able to identify limitations of all of the studies included in their summary on repeated oral readings and to provide information about the scarcity of experimental research on independent reading. Because this subgroup so carefully read the studies and noted their limitations, only a few details of individual studies are included in this review. It suffices to repeat their conclusion that repeated oral reading helps with word recognition (gain effect of .55), fluency (gain effect of .44), and comprehension (gain effect of .35) and repeat the limitations they provided: that there was no information about transfer for the Immediate Effects studies, that there was insufficient information to detail findings by grade level (Group Experiments studies), that the findings of the Single Subjects studies could have been magnified by the fact that all instruction was one-on-one with low-achieving students, and that there were no true control groups in the methods studies (pp. 3-18 and 3-19).

This leaves teachers in the position of needing to conduct their own research into the effect of repeated readings and to consider implications for how the effect of

repeated readings might transfer to their classrooms. It also positions teachers to counter mandates that call for particular types of repeated reading at particular grade levels. The NRP report offers no SBRR about the appropriateness of such experiences at particular grade levels.

What Can Be Learned from the Studies of Fluency Included in the NRP Report?
Table 4 provides a summary of the instructional implications of the fluency studies cited in the NRP report. Although Samuels et al. (2000) outline a number of limitations on the studies cited in this report, it is clear that policymakers could better help readers if they adopted the definition of reading put forth by this subgroup and re-envisioned fluency as the antithesis of word-by-word reading and as a means of helping children read in thoughtful units as a way of making meaning from the text.

Table 4. Instructional Implications from Fluency Studies as Cited in the NRP Report

If you want students to . . .	*You could . . .*
More fluently read a particular text	Have them reread that text (e.g., Neill, 1980; Sindelar et al., 1990)
Better understand a particular text	Have them reread that text (e.g., Sindelar et al., 1990)
	Read the story aloud to them first (e.g., Reitsma, 1988; Rose & Beattie, 1986; Smith, 1979)

Helpful Quotes about Fluency Instruction from the NRP Report
As teachers work out for themselves (and address with others) the instructional implications of the fluency studies cited by the NRP, these comments from the report might prove useful:

> (Round robin reading): “These procedures have been criticized as boring, anxiety provoking, disruptive of fluency, and wasteful of instructional time, and their use has been found to have little or no relationship to gains in reading achievement (Stallings, 1980).” (P. 3-11)
>
> The purpose [of fluency instruction] is to help students through oral reading practice and guidance to develop fluent reading habits that would allow them to read text more quickly, accurately and with appropriate expression and understanding. (P. 3-11)
>
> Not surprisingly, all 14 [Immediate Effects] studies reported demonstrable improvements from a first passage reading to a final passage reading with whatever measures were used. . . . Given the lack of transfer measures in [these] studies, the greater gains for low readers could be an artifact of the design

because these readers' initial performances would be relatively more deficient and would therefore be most amendable to improvement. (P. 3-15)

What inferences can be made from this set of studies? It certainly cannot infer [*sic*] that repeated reading or other guided repeated oral reading procedures would be effective in raising reading achievement. . . . However, the . . . improvements . . . suggest the possibility that such procedures could have transfer effects worth examining. (P. 3-16)

These studies [Group Experiments] as a collection have not provided sufficient data to allow for a sound analysis of the relative impact of repeated reading procedures on students at different grade levels. (P. 3-17)

There were not enough comparisons [in the Methods Comparison studies] of guided repeated oral reading procedures to allow for a systematic determination of best procedures. For the most part, the comparisons that were done resulted in no differences. (P. 3-19)

The pattern of findings for these [Single Subjects] studies is almost identical to what was reported in the meta-analysis . . . but all but one of these studies were conducted with a one-to-one teacher-student ratio and all were carried out with low-level—sometimes very low-level—readers, and either of these factors could magnify the effect. (P. 3-19)

[T]hese studies found no gains in reading due to encouraging students to read more. It is unclear whether this was the result of deficiencies in the instructional procedures themselves or to weaknesses and limitations in the study designs. . . . [I]t is not that studies have proven that this [independent reading] cannot work, only that it is yet unproven. . . . To really understand the implications . . . it is important to compare these routines against procedures in which students actually read less. (P. 3-27)

One word of caution can be drawn from a short-term study (Anderson, Wilkinson, & Mason, 1991) [a study not included in the fluency report] that found that too much attention to fluency issues within a reading lesson could detract from reading comprehension. It should be noted that in all of these studies, the fluency work was only part of the instruction that students received. In most cases, the fluency work was relatively brief (15 to 30 minutes per lesson), and students who received these lessons were still engaged in other reading activities including comprehension instruction. Guided repeated oral reading and repeated reading provide students with practice that substantially improves word recognition and fluency and—to a lesser extent—reading comprehension. They appear to do so, however, in the context of an overall reading program, not stand-alone interventions. (P. 3-20)

5: Vocabulary

Subgroup Members: Michael Kamil, Ph.D., Chair (professor in psychological studies); Gwenette Ferguson, M.Ed. (reading teacher); Norma Garza, C.P.A. (certified public accountant); Thomas Trabasso, Ph.D. (professor in psychology); and Joanna Williams, Ph.D. (professor in psychology and education)

The same subgroup (Kamil, Ferguson, Garza, Trabasso, & Williams, 2000) that looked at the reading comprehension research detailed in the next chapter also looked at vocabulary research. Within this field, they chose to examine data about the measurement of vocabulary and about vocabulary instruction. The authors did not, however, find any research on vocabulary measurement that met the NRP criteria (p. 4-3). The researchers therefore chose instead to look at the type of measures used by researchers who study vocabulary. They reported that such researchers predominantly use only experimenter-generated measures and only rarely use standardized measures. They concluded that standardized tests "did not seem to be sufficiently sensitive to vocabulary changes" and that, in the classroom, assessing vocabulary growth "would be best done with teacher-generated instruments" (p. 4-24).

In terms of vocabulary instruction, this Comprehension subgroup (Kamil et al., 2000) argued that vocabulary "occupies an important position in learning to read. As a learner begins to read, reading vocabulary encountered in texts is mapped onto the oral vocabulary the learner brings to the task" (p. 4-3). The authors explained that words encountered need to first be in the reader's oral vocabulary, and thus vocabulary was "an important middle ground in learning to read" (p. 4-3). As with the other subgroups, the members of this group searched databases and eliminated studies not consistent with NRP criteria. As they did with their review of the comprehension literature and unlike the decisions made by the Alphabetics subgroup, the Comprehension subgroup focused on vocabulary studies with normally achieving students whose first language was English. They eliminated all studies that "dealt exclusively with learning disabled or other special populations" (p. 4-16). The Comprehension subgroup ended up with 50 studies and decided these did not constitute a numerically large enough database for a meta-analysis.

They then divided these studies into 21 different methods (e.g., keyword method, semantic mapping, contextual analysis). Many studies, however, appeared in more than one category—for example, the same 6 studies appeared both under keyword method and semantic mapping (p. 4-33). While the authors also offered broader conceptual categories (explicit instruction, indirect instruction, multimedia methods, capacity methods, and association methods), they found that this categorization system also was not practical since many of the 50 studies fit into more than one category.

Because neither of their categorization systems clearly distinguished the studies, I decided to simply divide the studies into four groups—indirect learning, explicit instruction, restructuring the task, and context. Across all categories, I found that the research supported the findings of the subgroup. Those findings were listed by the authors:

1. Vocabulary should be taught both directly and indirectly.

2. Repetition and multiple exposures to vocabulary items are important.

3. Learning in rich contexts is valuable for vocabulary learning.

4. Vocabulary tasks should be restructured when necessary.

5. Vocabulary learning should entail active engagement in learning tasks.

6. Computer technology can be used to help teach vocabulary.

7. Vocabulary can be acquired through incidental learning.

8. How vocabulary is assessed and evaluated can have differential effects on instruction.

9. Dependence on a single vocabulary instruction method will not result in optimal learning. (P. 4-27)

The subgroup authors note, however, that there was very little research on vocabulary instruction in the early grades. The conclusions of this subgroup, then, are based primarily on research conducted in grades 3 to 8. The authors also point out that most of the vocabulary that students acquire does *not* come from instruction:

> Estimates of vocabulary size seem to suggest that there would never be sufficient classroom time to instruct students to the level of their acquired vocabulary. This implies that much of a student's vocabulary will have to be learned in the course of doing things other than explicit vocabulary learning**.** Students may well pick up vocabulary in contexts different from the formal learning of a classroom reading group. It may even be that the vocabulary acquired in this way is more memorable, given the role of motivation in its acquisition because the vocabulary acquired in this way may be far more useful. Repetition, richness of context, and motivation may add to the efficacy of incidental learning. (P. 4-26).

Some of the authors of the indirect studies provide additional information about this finding. Robbins and Ehri (1994) cite a study conducted by Beck, Perfetti, and McKeown (1982) that showed that in 19 weeks of vocabulary instruction, fourth graders learned 85 targeted words. Beck et al. estimated that fourth graders learned between 1,000 and 3,000 other words during this same time period. Similarly, Senechal and Cornell (1993) estimated that children learn an average of 5 words a day from the time they are 18 months old until they are 6 years old. These comments mirror the conclusion accepted broadly in the field, which is that children learn most of this vocabulary incidentally. Indeed, Nagy, Anderson, and Herman (1987) concluded that

average reading in grades 3 to 8 probably accounts for one-third of a child's annual vocabulary growth (p. 55).

A close look at a few of the studies in two categories will perhaps make clear the robustness of the findings across all the vocabulary studies investigated.

Indirect Learning Effects

Dickinson and Smith (1994) studied storybook reading in 25 preschool classrooms for low-income children. Of the participants, 62 percent were European Americans, 34 percent were African Americans, and 4 percent were Mexican Americans. In their study, Dickinson and Smith sought to understand if there were identifiable patterns of teacher-child interactions and if the ways teachers read to children impacted their literacy development. In each classroom, videotapes were made of the teachers reading aloud to students. These tapes were transcribed and each utterance coded. In addition to coding who talked (teacher or students), when the talk occurred (before, during, or after the reading), and whether the response was a request for information, a response to a request, or a spontaneous offer of information, Dickinson and Smith also coded utterances as cognitively challenging, less cognitively demanding, and management. The possible codes were combined as appropriate, and this resulted in a list of 21 mutually exclusive variables. The researchers subsequently looked for combinations of variables that would distinguish classrooms. They determined that there were three distinctive approaches: the co-constructive approach, the didactic-interactional approach, and the performance-oriented approach. The co-constructive approach was "characterized by considerable talk as books are read. . . . In classrooms where this approach is used, book reading is an occasion when children and teachers enjoy being together, with discussion of the book being integral to this enjoyment" (p. 116). In the didactic-interactional approach, there was "limited talk, group recall of familiar or highly predictable text, recall of recently read text, and a high proportion of talk dealing with organizational matters" (p. 116). In the performance-oriented approach, the "reading of books [is treated] as a performance that is to be enjoyed and interrupted only for important matters" (p. 116).

A year later, when the children were 5 years old, they were given a battery of tests of language and literacy development. The Peabody Picture Vocabulary Test (PPVT) was used as a measure of receptive language skills. To assess story comprehension, the children were read a book, *The Snowy Day* by Ezra Jack Keats (1976), and asked a series of probe questions. Dickinson and Smith (1994) found that the performance-oriented approach "more successfully supported vocabulary growth than the didactic interactional approach" (p. 117). They also found that they were able to predict children's subsequent PPVT and story comprehension scores based on the amount of analytical talk that children used during the read aloud. Indeed, over half the variance in the PPVT scores could be predicted in this way. The researchers argued that the analytical discussions "helped create a stronger conceptual base for children's vocabulary while they also provided occasions for use of low frequency words" (p. 117). They noted that in order for children to be fully engaged in analytical discussions, they "need to become part of a teacher-student dialogue by actively contributing or by attending to the responses of others" (p. 118). In closing, the researchers noted that the relationship between lower vocabulary development and didactic-interactional

classrooms may suggest that “a steady diet of books with predictable text may not be optimal” (p. 119).

In a second study with very young children, Senechal and Cornell (1993) looked at the impact of storybook reading on the vocabulary development of 80 predominantly white middle-class 4- and 5-year-olds. The researchers were also interested in determining the impact of reading conditions on vocabulary acquisition. The book that the children heard contained 10 words that Senchal and Cornell believed would not be typically known to young children. They believed, however, that children would know synonyms for those words. For example, they expected children not to know the word *infant*, which appeared in the story, but to know the word *baby*. Five reading conditions were used. In the first condition, children were asked “what and where questions when target words were introduced” (p. 364). In the second condition, children were encouraged to label the target items. In the third condition, recasting, the teacher introduced the word and then repeated the sentence with a synonym. In the fourth condition, word repetition, the teacher repeated the words. In the final condition, children listened and were not encouraged to participate.

The children were pretested on both receptive and expressive vocabulary, tested again immediately after the book was read, and then tested a week later. On the immediate post-test for receptive vocabulary, the 4- and 5-year-olds recognized the same number of words. A week later, the 5-year-olds remembered significantly more words than did the 4-year-olds. All children, however, remembered more words a week later than they forgot. There was no statistically significant impact on expressive vocabulary. There were also no differences among the children based on the conditions under which the book was read. The researchers concluded that a single book reading could increase receptive vocabulary for 4- and 5-year-olds, with a greater impact on the 5-year-olds.

In a third study conducted with young children, on two different occasions 2 to 4 days apart, Robbins and Ehri (1994) read a story containing 11 target words to individual kindergarten children who could not yet read (n = 33). Some of the target words appeared twice in the story, and some appeared once. After the first reading, the children were asked what they liked about the story. After the second reading, the children described what they liked in the story and were asked if they liked it a lot, a little, or not very much and then were given a multiple-choice vocabulary test. The test assessed their knowledge of 22 unfamiliar words, half of which had appeared in the story. Like Senchal and Cornell (1993), Robbins and Ehri were interested in how reading to young children contributed to vocabulary development. In earlier grades, it is hypothesized that young children do not learn many words from the texts they read because the texts predominantly contain words they already know. Read alouds then become a possible source of new vocabulary for young children.

The researchers found that children who did better on the pretest (PPVT) knew more new words on the post-test than did children who did not do as well on the pretest. All children, however, did better with recognizing words they had heard in the story as opposed to words they had not heard. Some of the words heard four times were learned by most students; others were not. The researchers hypothesized that hearing words four times in stories may be necessary but not sufficient for learning. On the basis of

their findings, the researchers concluded that “listening to stories was an effective means of expanding” the children’s word knowledge (p. 58).

These three studies demonstrate that young children learn new words as a result of being read to. The first study also suggests that one way children engage with the text (high analytic) may contribute positively to expressive language acquisition.

Direct Instruction

Beck et al. (1982); McKeown, Beck, Omanson, and Pople (1985); and Wixson (1986) were all interested in how explicit vocabulary instruction could improve comprehension. In Beck et al., students in one fourth-grade classroom (*n* = 27) served as the experimental group. On the basis of pretest scores, they were matched with 27 children in two other classrooms. Those children served as the control group. The children were low-SES students; about 70 percent of them were African Americans. In the experimental group, 104 words were taught and reinforced over a 5-day period. A subset of these words was included in subsequent lessons. Students were taught these words 16–22 times from October until March. The words taught once were referred to as “some” words, and the words revisited were referred to as “many” words. During the 5-day cycle, students were introduced to words as part of a set, such as words about people, and were provided with a variety of interactions with each word. For example, students were asked if the target word *accomplice* and the word *crook* go together and why. They were also asked to say “yay” or “boo” to indicate how they would feel about someone called an *accomplice.* Over a 5-day period, students spent 2.5 hours on 8 words. They encountered each word at least 10 times. A total of 104 words were studied in this way.

Students in the experimental group did significantly better on all post-tests. On the vocabulary subtest of the ITBS post-test, students in the experimental group had a raw score gain of 43 percent as compared to 17 percent for the control group. This translated into grade-level equivalents of 4.5 for the experimental group and 4.0 for the control group. On the vocabulary post-test, the students in the treatment group knew more of the target words but did not know significantly more of the words not taught as part of the study. On the semantic decision and sentence verification tasks, students in the treatment group were more quickly able to identify taught words than were students in the control group. On the story recall tasks, students in the treatment group recalled more information about stories containing words that had been repeated across instructional periods.

The researchers pointed out, however, that even though students in the treatment group knew more of the target words than did students in the control group, they did not know them all. Instead, they knew 77.6 percent of the “some” words and 86.55 percent of the “many” words. They concluded that “[a]cquiring word meanings to a high level is not an easy task, even with extensive instruction” (p. 518).

In McKeown et al. (1985), some of the same researchers tried to determine how extensive instruction needed to be. Focusing again on fourth-grade students, the researchers provided three forms of vocabulary instruction—traditional (words and definitions), rich (elaborated word meanings in diverse contexts), and extended (use beyond the classroom). In all three forms, some words were taught 7 times and other words were taught 4 times. On a multiple-choice post-test, the researchers found that

children in all treatment groups knew more target words than did the children in the control group. They also knew more high-frequency words than low-frequency words. There were no differences in the numbers of known words based on type of instruction.

On the semantic decision task (a task designed to assess rate of word comprehension), the students in the extended treatment group had significantly faster reaction times. Words taught more frequently were recognized faster than were words taught less frequently. On a task intended to assess students' ability to "interpret contexts containing instructed words" (p. 531), students in the rich and extended treatment groups outperformed their peers in the traditional and control groups. On a measure of story comprehension, the researchers found that traditional instruction did not lead to comprehension gains, although both rich and extended instruction did so. Children who received rich and extended instruction also used the words spontaneously outside of class.

The researchers concluded that instruction "that affects word definition knowledge will not necessarily affect comprehension" (p. 532) and that "more encounters yielded better results than fewer encounters for word knowledge, for fluent access to word meanings, for context interpretation and for story comprehension" (p. 532). Extended instruction was a factor only on timed measures.

In a third direct instruction study, on consecutive days Wixson (1986) provided 120 average and above-average fifth graders with two 30-minute vocabulary lessons and then gave the students 15 minutes to silently read a story in which the words appeared. Two different methods were used—dictionary and concept. Two different types of words were taught—those central to the story and those not central to the story. There was no control group. On the third day of the study, students were asked to recall the story and then were asked particular questions about it. They were also given two measures of vocabulary: definition and example.

Wixson (1986) found that students who were taught central words knew more central than noncentral words and better understood ideas related to the central words. The reverse was also true—students who were taught noncentral words knew more noncentral than central words and better understood ideas related to noncentral words. This pattern did not hold for a follow-up task in which students who learned both central and noncentral words did equally well on ideas related to central words.

What these three studies suggest is that when students are taught the meaning of words that subsequently appear in stories, they do better on story comprehension than do students who do not receive such instruction.

Across the remaining categories that I used (restructuring the task and context), the researchers found that various ways of restructuring the task contributed to vocabulary acquisition and that a combination of "definitional and contextual approaches worked better than either method used alone" (p. 4-23). (See the summary from the subgroup on pages 4-18 through 4-24.)

What Can Be Learned from the Studies of Vocabulary Included in the NRP Report?

Because I found that the vocabulary studies supported the list of instructional implications provided by the subgroup (p. 4-27) and shown near the beginning of this chapter, no additional table of findings is included in this section. It is clear from the

research included in the report that children learn vocabulary both directly and indirectly and that understanding the meaning of words in a text increases comprehension.

Helpful Quotes about Vocabulary Instruction from the NRP Report
As teachers work out for themselves (and address with others) the instructional implications of the vocabulary studies cited by the NRP, the following comments from the report might prove useful:

> Estimates of vocabulary size seem to suggest that there would never be sufficient classroom time to instruct students to the level of their acquired vocabulary. This implies that much of a student's vocabulary will have to be learned in the course of doing things other than explicit vocabulary learning**.** Students may well pick up vocabulary in contexts different from the formal learning of a classroom reading group. It may even be that the vocabulary acquired in this way is more memorable, given the role of motivation in its acquisition because the vocabulary acquired in this way may be far more useful. Repetition, richness of context, and motivation may also add to the efficacy of incidental learning. (P. 4-26)

> Various ability levels and age differences can significantly affect learning gains from vocabulary instruction methods. The studies underscore the need to consider carefully the different impacts that various vocabulary instruction techniques can have for students of different ages and abilities, and, accordingly, the importance of selecting appropriate methods. (P. 4-18)

> One trend that was strongly reflected in the database was that high frequency and multiple, repeated exposures to vocabulary material are important for learning gains. In accordance with this finding, a trend was also noted that extended and rich instruction . . . was superior to less comprehensive methods. (P. 4-22)

6: Comprehension

Subgroup Members: Michael Kamil, Ph.D., Chair (professor in psychological studies); Gwenette Ferguson, M.Ed. (reading teacher); Norma Garza, C.P.A. (certified public accountant); Thomas Trabasso, Ph.D. (professor in psychology); and Joanna Williams, Ph.D. (professor in psychology and education)

The authors of this report (Kamil et al., 2000) began by noting that comprehension has been a focus of reading instruction for only the last 30 years. This research began with Markman's (1977, 1981) research showing that children were not noticing inconsistencies in texts. Since that time, reading has come to be seen as an active process, as "intentional thinking during which meaning is constructed through interactions between text and reader" (p. 4-39; definition referenced to Durkin, 1993). Kamil et al. noted that "the content of meaning is influenced by the text and by the reader's prior knowledge that is brought to bear on it" (p. 4-39; referenced to Anderson & Pearson, 1984). They explained that "the bulk of instruction of text comprehension research . . . has been guided by the cognitive conceptualization of reading" (p. 4-39). They further noted that the emphasis has been on teaching strategies, although "readers normally acquire strategies for active comprehension informally" (p. 4-5).

To conduct an analysis of comprehension studies, Kamil et al. (2000) searched various databases and found 453 studies to which they added 28 more. They then applied the criteria established by the entire panel (published in a scientific journal, experimental design, random assignment to treatment—which they "relaxed" in several studies). This left 205 studies. The students involved in these studies were in grades 2 to 6**.** The subgroup next divided these 205 studies into 15 categories (e.g., comprehension monitoring, cooperative learning, graphic organizers). Kamil et al. then decided not to do a meta-analysis because "even the studies identified in the same instructional category used widely varying sets of methodologies and implementations" (p. 4-42). Instead, they read, summarized, and drew conclusions about the impact of particular interventions. Their conclusion was that comprehension monitoring, cooperative learning, graphic organizers such as story maps, question answering, question generation, and summarization appear to improve comprehension. The authors also looked at multiple-strategy instruction and concluded, "the evidence supports the use of combinations of reading strategies in natural learning situations" (p. 4-46).

Kamil et al. (2000) did not find support for research in the other aspects of reading comprehension (curriculum plus strategies, listening actively, mental imagery, mnemonic, prior knowledge, psycholinguistics, story structure, teacher preparation, and vocabulary/comprehension relationship). In these "other" categories, it was not the case that the research proved a particular strategy did not improve comprehension; rather, Kamil et al. found that there was not enough evidence to show that particular strategies *did* improve comprehension. Lack of evidence predominately comes from five concerns raised by this subgroup: wide range of methods assessed, limited or no measures of reading comprehension, very small number of studies, mixed results, and not a representative population.

Wide range of methods: Curriculum plus strategy instruction. Kamil et al. (2000) concluded that there were such "variation and complexity of curricula across these studies" (p. 4-73) that there was not ample evidence to support or refute the effectiveness of any particular curricular intervention.

Mixed results: Listening actively. The Comprehension subgroup located 4 studies in which students received "direct instruction on learning to listen to others . . . who read" (p. 4-75) while following along in the text. Students did better on experimenter tests in 2 of the 4 studies and on standardized tests in 2 of the 4 studies.

Mixed results and not a representative population: Story structure. In the 17 studies that Kamil et al. (2000) found on story structure, half of the studies conducted with fourth graders and all the studies with fifth and sixth graders were conducted with poor readers. There were 2 studies that involved normally achieving third graders and 2 with normally achieving fourth graders. Experimenter tests included recall, question answering, and identifying the elements of a story. Three studies used standardized tests. The results from these measures were mixed, with some successes with some groups and none with others.

Mixed results and a very small number of studies: Vocabulary instruction and relation to reading. The subgroup found 4 studies that examined the relationship between teaching vocabulary and reading. Only 2 of these assessed comprehension. In one study, students in the experimental group outperformed students in the control group; in the other study, they did not.

Very small number of studies and not a representative population: Teacher preparation. While Kamil et al. (2000) note that it is important to know "how and whether teachers can be effectively prepared" to teach reading strategies in naturalistic environments (p. 4-94), they found only 4 such studies, and all were conducted with struggling readers. The authors conclude that "this small set of studies" shows that teachers can learn to teach for strategies and that students benefit when this happens. They call, though, for more research in this area "with a wider range of readers."

Lack of comprehension measures. (1) Mental imagery. Seven studies were located. Students who participated in those studies were in grades 2 to 8. Findings showed that when students were asked to form an image associated with a sentence, they had better memory of the sentence imagined than did students not asked to form that image. (2) Mnemonic. In the 2 mnemonic studies, eighth graders were asked to use keywords as mnemonics. Students in the treatment groups had increased memory and recall. (3) Prior knowledge. There were several methods of activating prior knowledge (e.g., asking the students to think about the topic, teaching related information, prereading, predicting, making associations, and previewing). The most common measure was a memory task. As Kamil et al. (2000) summarized, "Recall was used in nine [of the 14] studies, question answering was used in three studies, and achievement in content area was used in two studies" (p. 4-85).

Very small number of studies: Psycholinguistic approaches. The Comprehension subgroup quoted the definition of psycholinguistics put forth by Harris and Hodges (1995, p. 197): psycholinguistics is "the interdisciplinary field of psychology and linguistics in which language behavior is examined. . . . [It] includes such areas of inquiry as language acquisition, conversational analysis, and the sequencing of themes and topics in discourse." Using this definition, Kamil et al. (2000) found only one study

that “trained readers on a psycholinguistic skill” (p. 4-85). In that third-grade study, the skill was understanding the referents of pronouns.

Teachers need to understand, however, that instruction focusing on the cue systems in language (semantics, syntax, and grapho-phonemics in Goodman and Burke’s [1972] terms, or meaning, structure, and visual in Clay’s [1985] terms) are considered by many to be a psycholinguistic approach. From such a perspective, learning to read or developing as a reader is a psycholinguistic process because the focus is the acquisition of written language, or reading. Studies, then, that examined reading comprehension in whole language, literature-based, or Reading Recovery settings would have been considered by those others to be psycholinguistic studies. Such studies, however, were not included in the review of the comprehension literature conducted by this subgroup.

Studies in seven other categories were considered by Kamil et al. (2000) to be successful. A summary of the studies in these areas and the conclusions reported by the NRP are shown below.

- Comprehension monitoring (22 studies, grades 2 to 6) helps students with tasks that require them to monitor comprehension.

- Cooperative learning (10 studies, grades 3 to 6) helps with reading comprehension strategies and performance on reading tests.

- Graphic organizers (11 studies, grades 4 to 6) help students remember information. Children who already have skills in reading and writing benefit most.

- Question answering (17 studies, grades 3 to 5) helps students score higher on “experimenter tests of question answering” (p. 4-45).

- Question generation (27 studies, grades 3 to 9), helps with generating questions and answering inferential questions. However, “only 3 of 13 effects were statistically significant, casting doubt on the generality of this single strategy instruction” (p. 4-45).

- Summarization (18 studies, grades 5 to 6) helps with memory and identifying main idea.

Agreeing with Rosenshine and Meister (1994), Kamil et al. (2000) noted, however, that not enough was known about instruction because “not enough studies have been devoted to implementation” (p. 4-47). They also pointed out:

> Effective reading instruction is associated more with independent teacher action than with implementation of basal text prescriptions. . . . Developing metacognitive readers who understand their reasoning requires teachers who themselves understand their reasoning, as well as a supportive environment in the schools for strategy learning. (P. 4-48)

A brief overview of some of the studies in these categories will perhaps give a sense of the research in these seven areas. K–2 teachers need to keep in mind, however, that none of the studies in the categories deemed to be successful looked at reading comprehension of kindergartners or first graders, and only 3 studies involved second graders. This is important for all teachers—and particularly important for Reading First teachers who work with children in K–3—who are being told to rely for guidance on the NRP report. The simple fact is this: There is no information in the NRP report about reading strategies that can be used to improve comprehension instruction for children in kindergarten and first grade. There is very limited information about reading strategy instruction for second-grade children. Teachers need to rely on other sources of SBRR to help young children better comprehend text.

Comprehension Monitoring

The Comprehension subgroup (Kamil et al., 2000) cited the definition of comprehension monitoring provided by Harris and Hodges (1995, p. 39): “Comprehension monitoring in the act of reading is the noting of one’s successes and failures in developing or attaining meaning, usually with reference to an emerging conception of the meaning of the text as a whole, and adjusting one’s reading processes according[ly]” (p. 4-69). This is much simpler in classroom life than it is in researcher-generated definitions. What the authors were arguing is that students need to be making sense of texts when they read. The students need to notice when the text does not make sense and should have a variety of ways of stopping to make sense of the text (they can reread, read ahead and then come back, think about what makes sense, look at the picture for clues, look again at the word, etc.). It is therefore logical, I think, to most teachers that all readers need to monitor comprehension in order to understand what they are reading. What this body of research examines is how to teach children to do that monitoring.

In Chapter 7 of this guide, I suggest that teachers need to think about how their beliefs align with the beliefs of the members of the various subgroups. I give the example of PA and note that the members of the Alphabetics subgroup (Ehri et al., 2000) seemed to assume that PA needed to be taught, whereas other researchers have argued that we ought to be focusing on how most children acquire this awareness independent of instruction. Something similar applies here. While most of the field agrees that all of us need to be monitoring for meaning while we read (just as most of us agree that children ought to know that language is made up of parts that can be blended and segmented), there are differences about whether all children need to learn this through instruction. One group argues that monitoring comprehension is a developmental process and that children need direct instruction in order to effectively monitor. Another group argues that because most children grow up in a meaningful world, they expect texts, like everything else, to be meaningful. There are differences of opinion, then, about who needs comprehension instruction: all children or simply those who do not understand that text is supposed to make sense and thus often read word-by-word, trying to get each word right and not holding onto meaning. It is also important to keep in mind that sometimes word-by-word reading is the result of the lack of attention paid to comprehension/meaning in kindergarten, first-grade, and sometimes second-grade reading instruction. As one second-grade child explained to me in a

clinical setting several years ago, “Do you want to know about real reading or reading in school? In real reading, there are stories. When we read in school, we just read words.”

In their search for studies that *taught* children to self-monitor, Kamil et al. (2000) found 20 studies. Three of those were with second graders, 6 with third graders, 8 with fourth graders, 5 with fifth graders, and 6 with sixth graders. In their list of comprehension monitoring studies, however, they list 22 studies. Since the Markman (1977) study was not an attempt to teach comprehension monitoring, I assumed it was cited as a reference but not as an included study. I was not able to ascertain which other study was cited but not included. The authors report, however, that all 20 studies were attempts to teach children to pay attention to meaning while they listened or read. In all studies, according to Kamil et al., the children in the treatment group were better able to do this than were children in the control groups. I was already familiar with the research in this area; for this review of those studies, I reread 12 of the 22 studies listed in Appendix A under Comprehension Monitoring. A brief overview of 4 studies provides a snapshot of the comprehension monitoring research cited in the NRP report.

Baumann, Seifer-Kessell, and Jones (1992) provided two forms of instruction—Think Aloud (TA) and Directed Reading Thinking Activity (DRTA)—and compared the impact of both to children in a Directed Reading Activity (DRA) control group. Sixty-six fourth graders were assigned to one of these three groups. Both forms of intervention were intended to improve students’ comprehension monitoring. Students were given pretests on the first day, received 10 days of instruction, and then were given post-tests for 3 days. The TA lessons consisted of helping students question as they read; consider both themselves and the text as sources of information; stop and ask, “Does this make sense?”; predict, read, and verify; infer; retell; reread; and read on. DRTA consisted of having students repeatedly make and confirm predictions. DRA instruction, the control, involved introducing new vocabulary, activating background knowledge, setting purposes, asking comprehension questions, and discussing the selection. On three of the measures used in the study (error detection text, comprehension monitoring questionnaire, and Degrees of Reading Power standardized test), Baumann et al. found that students in both treatment groups outperformed students in the control group. The researchers suggested that more research was needed to determine which method was more effective.

In the Block (1993) study, research assistants provided 178 children in 8 classrooms with strategy lessons for 90 minutes a day, 2 days a week, for 16 weeks. Another 174 students in 8 other classrooms served as the control group. The students were ethnically diverse (68 percent Mexican American, 16 percent African American, 14 percent European American, and 2 percent Asian American) and in grades 2 to 6. In the strategy lessons, children were introduced to a thinking and reading comprehension strategy and then to selected children’s literature and were asked to use that strategy as they read. The strategies included basic cognitive operations, fundamental thinking processes, decision-making strategies, problem-solving strategies, metacognitive strategies, creative thinking strategies, strategies for working effectively in small groups, and strategies for studying and working more effectively alone (p. 141). The students in the experimental group scored higher on the reading comprehension and vocabulary sections of the ITBS and, according to a post-test writing sample and interview question, were “better able to transfer their reading and thinking strategies to situations

outside of school" (p. 145). Block was also interested in students' self-esteem, social competence, idea generation, reflective thinking, problem solving, and ability to generate alternatives. She found that the students in the experimental group outscored the control group on all measures designed to assess these factors.

In the Tregaskes and Daines (1989) study, a cloze test and an error detection test were used as pretests and post-tests to assess reading comprehension of two groups of low-SES European and Mexican American sixth-grade students. Half of the 152 students were in three treatment classrooms, and half were in three classrooms that served as the control group. Teachers of students in the experimental classrooms were trained in five meta-cognitive strategies: summary sentences, visual imagery, webbing, self-interrogation, and click-cards (a self-monitoring system). They then chose which strategies they wanted to use and decided how "frequently and for how long their students would practice the skills" (p. 57). Strategy instruction was provided during social studies and science periods. At the end of 12 weeks, the students in the treatment classrooms had significantly higher scores on both the cloze test and the error detection test.

Paris, Cross, and Lipson (1984) conducted two sets of studies inquiring into comprehension monitoring. In the first study, they provided instruction in informed strategies for learning (ISL) to two groups of third graders and two groups of fifth graders for 30 minutes a day, 2 days a week, for 4 months. Students in two other classrooms at each grade level served as the control groups. When the treatment groups were receiving the ISL instruction, the other two groups received "non-related" instruction. Paris et al. reported that while the comprehension scores of the treatment and control groups were not significantly different, the children in the treatment group were "more aware about reading" (p. 111), better understood the usefulness of comprehension-monitoring strategies (e.g., plan ahead, look back, ask for help), and performed significantly better on cloze and error detection tasks. In the second phase of their study, the researchers taught 50 teachers how to provide an expanded version of ISL as a part of their regular reading instruction. There were also 25 classrooms used as the control classrooms. The children in the treatment groups again outscored their control-group peers on all measures except for standardized reading tests.

Across the comprehension-monitoring studies, then, it becomes clear that children in grades 3 to 6 can be taught to pay attention to meaning when reading and to use a variety of strategies when the text is not making sense. It is also clear that children who have been taught to monitor for meaning and use comprehension strategies do better on various reading assessments than do children who have not received such instruction**.** Kamil et al. (2000) cautioned, however, that "the number of studies . . . is small [and] the method does not seem to generalize for 2nd graders" (p. 4-44). They conclude that, nevertheless, "it may be a useful addition to a program that employs flexibility and the teaching of multiple comprehension strategies" (p. 4-44).

Cooperative Learning

The evidence for cooperative learning as an effective comprehension strategy was not as straightforward. The members of this subgroup located 10 studies (9 written in English) on cooperative learning, 2 each at grades 3 to 6. On the basis of their review of these 10 studies, the researchers concluded that "having peers instruct or interact over

the use of reading strategies leads to an increase in the learning of strategies, promotes intellectual discussion, and increases reading comprehension" (p. 4-71). On the basis of this summary, I expected that the studies cited here would provide the same strategy or skill instruction to both groups and that the only difference between groups would be that in the treatment group students worked together, whereas in the control group students worked alone. I also expected that most or all studies would include comprehension measures. This, however, turned out not to be the case.

Klingner, Vaughn, and Schumm (1998), Judy, Alexander, Kulikowich, and Wilson (1988), and Guthrie et al. (1996) are examples of studies that did not meet my expectations. Klingner, Vaughn, and Schumm began their article by noting the need for the development of "effective, empirically based interventions that seem viable to teachers" (p. 4) and that would allow teachers to balance a "wide range of students [*sic*] needs in heterogeneous classrooms while facilitating the acquisition of reading skills and content-area knowledge" (p. 4). To this end, the researchers divided a group of fourth-grade students into treatment and control groups. There were three classrooms (85 students) that received the treatment and two classes (56 students) that served as the control group. Students in both groups learned the same content and were together for 11 instructional sessions of 45 minutes each. Students in the treatment group learned Collaborative Strategic Reading, which consisted of learning how to preview a passage, monitor it for comprehension, get the gist of it, and wrap up. After teachers introduced this information, students tried out the strategy and "were supported in their efforts to do so" (p. 7). As students became more proficient, they modeled this strategy for others, and from the fourth day on, students worked together in heterogeneous groups. In the control group, teachers followed directions in the manual for the history textbook used. These directions included having the teacher introduce the vocabulary, having the students look at pictures, reading the text together, and having the teacher summarize, ask questions, and lead a discussion. One group then received specific reading process instruction (Collaborative Strategic Reading) and the other group did not. The first group had the opportunity to collaborate; the second group did not. Although assessment measures favored the treatment group, it is therefore not possible to say whether it was the strategy instruction itself, strategy instruction combined with cooperative learning, or cooperative learning itself that resulted in the reported gains.

Guthrie et al. (1996) used Concept Oriented Reading Instruction (CORI) as an instructional framework with third-grade and fifth-grade students. There are collaborative aspects to CORI. However, there was no control group. No conclusions can then be drawn about the impact of the collaborative component of CORI.

Judy et al. (1988) were interested in how two different instructional approaches improved the analogy performance of gifted and nongifted sixth graders. A total of 194 students participated in the study; two-thirds of them were in the treatment groups. Those students received either direct or inquiry-based instruction. Following instruction, 24 students were selected to be peer tutors and taught analogical reasoning to another sixth grader. Conclusions were drawn about the effectiveness of the treatment method: the direct instruction method led to higher scores on the analogy tasks. Conclusions were also drawn about peer tutoring: the students who were tutors did no better on the analogy tasks than did the students who were not tutors. The students who were tutored by peers scored about the same on the analogy task as students who were taught by

adult teachers. No conclusions, then, can be drawn about the impact of analogy training on reading comprehension or on the success (or not) of peer tutoring.

Stevens, Slavin, and Farnish (1991) is the one study in this category that did meet my expectations; however, there were no differences between the reading performances of the students in the treatment group and the students in the control group. In this study, the researchers were interested in how direct instruction in strategies for identifying the main idea would impact students' abilities to identify the main idea. They also wished to understand how cooperative learning would impact students' learning. The authors had previously conducted research on the Cooperative Integrated Reading and Composition (CIRC) program (Stevens, Madden, Slavin, & Farnish, 1987). CIRC is designed for use by students in classrooms in which teachers meet with small reading groups and provide follow-up work to students after that meeting. In the small group, teachers introduce vocabulary, set a purpose, and discuss the story with the students after they have read it. Follow-up work includes reading the story silently, discussing and answering questions about the story, practicing vocabulary, practicing writing sentences using new vocabulary, practicing new spelling words, and writing a paragraph or short composition on a topic related to the story (p. 11). In addition to these basal-related activities, one day a week teachers provide direct instruction on a specific reading comprehension strategy such as drawing conclusions. These strategies are then practiced as part of follow-up activities. A third component of CIRC involves students in writing using a process approach.

The research on CIRC found that students who participated in CIRC had higher achievement test scores and higher scores on a writing test than did students in the control groups. These researchers, however, wanted to understand the relative impact of each of CIRC's components. To do so, they divided 486 ethnically diverse third and fourth graders into treatment and control groups. The two treatment groups were Direct Instruction with Cooperative Learning (CL) and Direct Instruction in Reading Comprehension (DI). The third graders had reading instruction for 90 minutes a week, and the fourth graders had reading instruction for 60 minutes a week.

For 4 weeks, in the CL group, students spent half their instructional time learning the main idea within a CIRC framework. During the other half of the time, students completed follow-up activities in small groups, working together to reach consensus on answers. Once they were consistently answering questions correctly, they worked independently. In pairs, they assessed each other's work. If a student made one error, he or she again worked independently and had a peer assess his or her work. If the student made more than one error, the teacher provided assistance. If the student made no errors, she or he took a mastery test. In the DI treatment, students also learned about the main idea within the CIRC framework, worked independently with teacher feedback as needed, and then took a mastery test. In the control group, teachers used traditional methods.

Students in both treatment groups outperformed students in the control group on main idea questions. There was no statistical difference between the CL group and the DI group on that measure. There were no differences among the three groups on inferential questions. This study, then, does not provide support for cooperative learning as a means of improving reading comprehension.

Two other studies included under the subgroup's category Cooperative Learning also investigated CIRC, although neither of those produced findings that supported cooperative learning. In the Stevens et al. (1987) study, no comparison was made between groups receiving similar instruction—one group that worked independently versus a second group that did not. In the Bramlett (1994) study, the treatment group worked within the CIRC framework, but the control group did not. Therefore, no comparisons could be made between CIRC with and without cooperative learning.

Kamil et al. (2000) included 4 other studies under the category Cooperative Learning. Three of those I could not find; one was written in Spanish, which I did not try to find as I am illiterate in Spanish. Of the 6 studies I did read, however, only one provided two groups of students with instruction that was identical except for the presence or absence of cooperative groups, and in that study, there were no differences between the treatment groups on main idea or inferential questions. Despite the finding reported by the NRP, then, I would argue that the studies cited in the NRP report provide insufficient data to support a claim that cooperative learning impacts reading achievement**.** I would suggest that cooperative learning be moved from the list of successful strategies to the list of strategies for which there is not sufficient evidence.

Graphic Organizers

Kamil et al. (2000) identified 11 studies that investigated the use of graphic organizers. One study involved second graders, and a second involved third graders. There were two studies each conducted with seventh graders and eighth graders. The bulk of the studies were with fourth graders (*n* = 5), fifth graders (*n* = 4), and sixth graders (*n* = 6). Kamil et al. concluded that graphic organizers help with memory of what was read and also help with comprehension. In 2 of the studies, the researchers reported gains on standardized achievement tests, whereas 7 used recall of content and 3 reported gains in content-area knowledge. I located and read 9 of the 11 studies included in this section and concur with this finding. The 2 studies described below are representative of the research in this area.

Alvermann and Boothby (1986) provided graphic organizer instruction during 25 minutes a day for 14 days, for 7 days, or for no days to 24 fourth graders with average or above-average reading ability. As a post-test, students were asked to recall what they read, and independent judges, using a template, rated the students' recalls. The researchers found that students who received 14 days of instruction were able to remember more information from an unrelated text than were the students who received either 7 or no days of instruction.

Darch, Carnine, and Kameenui (1986) asked 84 sixth-grade students to learn content-area information for 45 minutes a day for 15 days, in one of four ways: graphic organizers and group work (GO/Group); graphic organizers and individual work (GO/Individual); an approach called SQ3R (which involved only individual work), and Directed Reading/Group, which served as the control. The post-test was a series of questions (short answer or fill-in-the-blank) tied to the topic the students had studied. On the post-test, the students who learned via graphic organizers and who worked with each other had significantly higher scores on the post-test than did any of the other students. The students in both GO groups outperformed students in the other two groups. On a transfer test, however, while the GO groups again outperformed the other

students, there was no difference between the GO students who worked in a group versus those who worked alone. This study, then, could have been placed under the subgroup's Cooperative Learning category. However, because of mixed results for the GO/Group versus GO/Individual, it does not provide support for the use of cooperative learning.

Question Answering

The two "successful" approaches reviewed above, comprehension monitoring and graphic organizers, both help children improve their comprehension of texts. As an approach, question answering has a more narrow focus. The intent of this body of research is to determine whether teaching students how to answer questions helps them answer questions about passages. Kamil et al. (2000) placed 17 studies in this category and commented that "the effects of this method, however, are small" (p. 4-86). Consider, for example, the series of studies conducted by Raphael and colleagues (Raphael & Pearson, 1985; Raphael & McKinney, 1983; Raphael & Wonnacott, 1985). In these three studies, Raphael and her colleagues taught sixth graders, fifth and eighth graders, and fourth graders (respectively) to identify the type of question they were being asked. The question types came from a taxonomy developed by Pearson and Johnson (1978). This taxonomy identified three types of question-answer relationships (QARs): text explicit, text implicit, and script implicit. Script implicit are questions that are connected to a reader's schema or script. In all these studies, Raphael and colleagues explored how teaching about QARs could help students better identify where they would find the answers to text-based questions. Across all three studies, with the exception of what Raphael and Pearson (1985) consider a spurious result for eighth-grade students, the students in the treatment groups were better able to answer text-based questions than were their peers in control groups. In the delayed measure for fourth graders, there was no difference between the treatment and control groups, but there was a statistically significant difference on a near transfer measure.

Seretny and Dean (1986) randomly divided 54 second graders into two classrooms. The students were then assigned to three reading groups—high, medium, and low. Half of the students in each group became the treatment group, and half became the control group. In both groups, the students read a story a week from the basal reader. In the treatment group, teachers asked the students questions after each page was read orally by the students. In the control groups, students read orally but no questions were asked during the reading. At the end of 4 weeks, all students were given the Science Research Associates Achievement Test Battery. There were no differences in the reading scores of the above-average readers. There were, however, significant differences between the scores of the average and below-average readers in the treatment group versus the scores of those same types of readers in the control group.

In the Garner, Chou Hare, Alexander, Haynes, and Winograd (1984) study, the researchers' approach to increase correct responses on comprehension questions was to teach a group of 12 remedial readers (ages 9 to 13) to look back in the text to find answers. The researchers also taught the students why, when, and where such look-backs would be appropriate (i.e., look-backs should be used because one cannot remember everything one reads; they should be used when the question refers to the text and skimming the entire text to find a key segment is appropriate). Training took

place over 3 days and lasted between 15 and 20 minutes a day. Why, when, and where were introduced on the first, second, and third day, respectively. Six passages from the Reading for Concepts series (Liddle, 1977) were used both in the treatment and for the assessment measure. Students in the control group worked individually or in pairs on the same texts using text-processing strategies other than look-back. Five days after training, students were asked to read two unfamiliar texts and answer three questions. Students in both the treatment group and the control group got about the same percentage correct on answers they recalled (31.3 percent for treatment versus 33.3 percent for control). The students in the treatment group spontaneously looked back more often (13 percent cued by the administrator versus 100 percent cued for the control) and had significantly more correct answers (69.5 percent for treatment versus 22.3 percent for the control). In their discussion section, the researchers raise the question of why look-back instruction is needed and suggest that it seems that many upper elementary and middle school students consider it “illegal” to look back to the text. They note:

> Evidence of this perception was abundant within the study. During training, and even in testing after much prodding to re-access text, subjects reacted with surprise to the idea of using look-backs; they asked repeatedly for verification of the possibility. (P. 796)

The researchers added that during the testing of the control students, many of them complained that they couldn’t remember answers, but none of them spontaneously looked back at the text. The authors concluded, “Just as school is a likely candidate for the place where this conception originated, it is an excellent candidate for the spot where it can be dismantled” (p. 796). They suggest that teachers change from read-and-remember routines to read-remember-reread routines.

These and other studies included under the subgroup’s Question Answering category suggest that there are some students who do not spontaneously employ strategies that we, their teachers, use “naturally”—such as thinking about where we might find the answers (Raphael & Pearson, 1985; Raphael & McKinney, 1983; Raphael & Wonnacott, 1985), thinking about what is happening as we proceed through the story (Seretny & Dean, 1986), and looking back in the text to find answers to questions (Garner et al., 1984**).** Helping students learn to do these things increases their ability to answer questions about the text. The implication from all the studies in this section is that teachers ought to be helping students learn to do what they themselves do “naturally” as good readers.

Question Generation

Kamil et al. (2000) reported that the “strongest scientific evidence was found for the effectiveness of asking readers to generate questions during reading” (p. 4-45). Participating students were third to sixth graders. It is not clear from the Kamil et al. subgroup report how the 27 studies included in this category were determined. The authors noted that they relied on a review conducted by Rosenshine, Meister, and Chapman (1996). In that review, Rosenshine et al. identified 30 studies that focused on generating questions either singly or “in combination with other reading strategies” (p. 4-

88). However, in their appendix of studies included, the authors list not the 30 studies reviewed by Rosenshine et al. but, without offering an explanation of why 5 studies were excluded, only 25 of those 30. Two other studies not included in the Rosenshine et al. review are also listed among those included in the analysis. This total, 27, matches the total number reported on page 4-45 but is less than the 30 reported as being included in Rosenshine et al. It is also important to note that among the 25 attributed to Rosenshine et al., there are 7 unpublished doctoral dissertations and 2 papers presented at national conferences. Without these additional studies, there would have been 16 studies from Rosenshine et al. in this category.

The authors of the fluency report included the unpublished dissertations and conference papers in the 25 articles they included from Rosenshine et al. (1996). They organized those 25 studies into 6 subcategories: signal word prompts (6 published studies, 2 unpublished dissertations, 1 paper); generic questions or question stem prompts (3 studies all by the same author about college students attending lectures), main idea prompts (3 published studies, 2 dissertations); question-type prompts (1 published study, 1 dissertation, 1 paper), story grammar prompts (2 studies), and no prompts (1 study, 2 unpublished dissertations). The additional 2 studies were not categorized.

In order to be consistent with the decision of the NRP panel to include only studies published in refereed journals, for my review, I excluded the unpublished dissertations and conference presentations. I located and reviewed 13 of the 16 studies, all of which were published in refereed journals. The following studies are representative of the research conducted on question generation.

Davey and McBride (1986) divided 125 sixth-grade students into four treatment groups and one control group. All groups read the same passages—passages in which the main ideas were not explicitly stated. The question-training group was taught to generate questions that linked information across sentences and that were tied to the most important information. Two question-practice groups read passages and answered four open-ended questions. One of these groups answered literal questions, and the other answered inferential questions. The question-generation-practice group read passages and generated two questions. The no-question control group read the same passages but completed a vocabulary exercise instead of working with questions. The experimenter-designed measure asked students to read passages, generate two questions, and answer four literal and four inferential questions. The question-training and question-practice groups outperformed the inference-practice group and the no-question control group on literal questions. There was no statistically significant difference between the question-training and question-practice groups and the literal practice group. On the inferential questions, the question-training group outperformed all other groups. In other words, students taught how to make inferences did better on inferential questions than those who were not taught this.

Nolte and Singer (1985) were interested in whether third and fourth graders could use their understanding of story structure to ask their own questions and whether such questions would lead to improved comprehension. On the first of 10 days, the teacher modeled this practice for about 20 minutes, and then the students practiced asking questions. Over the 10 school days, the teacher modeled less and the students practiced more. In the control group, teachers wrote difficult words on the board for

students, asked them to read silently, and then gave a comprehension test. At the end of 10 days, both groups of students were asked to read a story and complete a 20-item test. Students in the treatment group averaged 12.28 correct answers, and students in the control group averaged 9.9 correct. This difference was statistically significant.

Lysynchuk, Pressley, and Vye (1990) conducted research on Palinscar and Brown's (1984) Reciprocal Teaching approach. As part of this instructional method designed to help students who are good decoders but weak comprehenders, students are scaffolded by their teacher into learning how to question, summarize, clarify, and predict. In the study, 72 students in grades 4 and 7 considered to be good decoders but poor comprehenders were placed either in a Reciprocal Teaching treatment group or a control group in which the teacher provided help with decoding and vocabulary as needed. The experiment lasted for 13 days, including 2 days of independent practice. The students in the Reciprocal Teaching group had significantly higher standardized test gains than did students in the control group. The experimental group gained a significant 9.97 percentile points while the control group gained a nonsignificant 1.63 percentile points. The students in the treatment group also outperformed the control group on daily retellings and comprehension questions.

Taylor and Frye (1992) also conducted research on Reciprocal Teaching. For 4 months, fifth-grade and sixth-grade students were taught Reciprocal Teaching strategies using their social studies texts. Two fifth-grade classrooms and two sixth-grade classrooms served as treatment groups; four others (two at each grade level) served as controls. No teacher volunteered to have his or her lower-level readers participate, so this study included only average and above-average readers. Taylor and Frye found that students in the treatment groups became better at summarizing materials. They were, however, not better at generating questions. The results for scores on short-answer measures were mixed.

Multiple Strategies

The authors of the Comprehension report were also interested in the impact of treatments that focused on multiple strategies. They considered Reciprocal Teaching to be a multiple strategy method; for that reason, there are references on Reciprocal Teaching under the Question Generation and Multiple Strategies categories. In their review of multiple strategies literature, the Comprehension study group again relied on a review by Rosenshine, this time a review he and Meister conducted in 1994. The reference pages for this subgroup (pp. 4-61 through 4-63) list 14 studies from Rosenshine and Meister's review of Reciprocal Teaching, an additional 12 studies on Reciprocal Teaching found by the subgroup, and 14 other multiple strategies studies selected by the subgroup. Excluding the dissertations, papers, and an article written in Spanish, the 14 studies from Rosenshine and Meister also were included in the subgroup's review of question generation studies. There are two errors in the list of additional studies, as 2 studies listed there are also included in the preceding Rosenshine and Meister list. That leaves 9 Reciprocal Teaching studies not included under the Question Generation category. The findings from those studies are consistent with the findings reported by Lysynchuk et al. (1990) above: Reciprocal Teaching helps students in grades 4 to 7 improve their scores on comprehension measures.

I read 11 of the 14 other multiple strategy studies selected by the subgroup. It was difficult to draw conclusions broadly across these studies because there was such a wide variety of strategies taught and outcome measures used. Indeed, the only overlap was that 2 studies were attempts to help students do better on inferencing questions. Carr, Bigler, and Morningstar (1991), for example, used cloze procedures, self-monitoring checklists, or structured overviews with sixth graders as means of improving students' abilities to make inferences. Students in one treatment group used the cloze and a checklist; students in the other used all three methods. The students in both control groups were better able to answer inferential questions. Students whose reading abilities were below average profited more than did average or above average students. To the same end, Reutzel and Hollingsworth (1991a) highlighted key vocabulary and used a generative reciprocal inference procedure with third graders and found that the reciprocal inference procedure led to higher scores on inferencing questions.

The other studies had little to do with each other. Adams, Carnine, and Gersten (1982), for example, taught study skills to fifth graders who had adequate decoding skills and found that the 15 students in the treatment group did better on a short answer measure, but there was no difference between the two groups on the retelling measure. The researchers also found that the students in the treatment group did not consistently use on the post-test measures what they were taught. Reutzel and Hollingsworth (1991b) divided sixth-grade students into four groups—one each on locating details, drawing conclusions, finding the sequence, and determining the main idea. The Barnell Loft Specific Skills Test was used before and after, and there were no differences among the groups on either the pretest or the post-test. Smith, Johnson, and Johnson (1981) were interested in knowing if students retained more information when it was presented in a controversial format (students argued pro and con) versus a concurrence approach (they were to avoid arguing) versus an individualistic approach (students learned on their own). They found that the sixth graders in the controversy group scored higher on several measures including achievement. Stevens et al. (1987) compared their CIRC model to traditional instruction and found that the third and fourth graders in the experimental group were better able to answer questions about main ideas. There were no differences between treatment and control scores on inferential questions.

Two studies dealt with adolescents: Anderson and Roit (1993) focused on severely reading-delayed adolescents, and Stevens (1988) focused on students in grades 6 to 11 who were reading 2 or more years below grade level. Since there are only two of these studies with adolescents and two different methods, no conclusions can be drawn about multiple strategy use with adolescents.

One study stood out from not only the rest of the multiple strategies studies but also the rest of the studies cited by most of the subgroups. It stood out because of the length of the study (across an academic year) and the depth of understanding of the teachers involved. The Brown, Pressley, Van Meter, and Schuder (1996) study looked at the strategy knowledge of students taught strategies by their teachers across the year versus the strategy knowledge of students not taught strategies. Rather than explain a procedure to teachers and then implement it for a limited number of days (or have outside others teach the children), the researchers identified five second-grade teachers involved in a project called Students Acquiring Independent Learning (SAIL).

SAIL teachers used transactional strategy instruction (TSI) in their classrooms. TSI is defined as the "joint construction of reasonable interpretations by group members as they apply strategies to texts" (p. 19). The long-term goal of TSI is the "internalization and consistently adaptive use of strategic processing whenever students encounter demanding text" (p. 19). The goal of the Brown et al. study was to compare the knowledge base of the students of SAIL teachers with the students of five other teachers, all considered exemplary, but who did not offer strategy-based instruction. All participating teachers worked in the same district with similar populations. Brown et al. chose this design because they felt that existing research on single-strategy instruction was limited to what students could do under research conditions and that "[s]eldom was generalized use of individual instructed strategies observed, nor was there evidence of generalized improvement in reading" (p. 18).

Sixty low-achieving second-grade students participated in this study. They were interviewed and tested in October and November, when strategy instruction was first being introduced, and then again in March and April. Brown et al. (1996) found that SAIL students had greater strategy awareness and use, greater acquisition of information from material reading in their reading group, and superior performance on standardized reading tests.

Kamil et al. (2000) concluded across these studies that "the evidence supports the use of combinations of reading strategies in natural learning situations" (p. 4-46). I misunderstood this statement the first few times I read it and now think that the first part of the sentence means that the research suggests not only that single-strategy instruction can help students better comprehend, but so can multiple-strategy instruction. The only multiple-strategy instruction approach, however, that has substantially been investigated in the literature is Palinscar and Brown's (1984) Reciprocal Teaching method. There are simply too few studies (sometimes only one study) conducted with any of the other multiple-strategy measures.

Relative to the subgroup's use of the prepositional phrase "in natural learning situations" (p. 4-46), I can only assume this relates to Brown et al. (1996) as it is the only study I found that could be considered to have occurred under natural, rather than contrived, settings.

If this is their meaning, then I concur. The studies in this category support the claim that students can be taught multiple comprehension strategies at the same time, and when that happens, scores on particular types of assessments increase. Across studies, this has consistently been the finding for one multiple comprehension method, Reciprocal Teaching (Palinscar & Brown, 1984).

Summarization

These 18 studies used a variety of methods to help students recognize and summarize main ideas. The students in these studies were predominately in grades 5 and 6. To assess effect of treatment, students were asked to recall, answer questions, or write summaries. Only 2 studies used standardized tests. The following 2 studies are representative.

Carnine, Kameenui, and Woolfson (1982) asked a group of middle-class fifth graders to read three stories and to answer inference questions. Students who read accurately and fluently but who could not correctly answer two of the three inference

questions qualified for this study. The first 36 students who fit the pattern were selected to participate. They were randomly divided into three groups: systematic instruction, corrective feedback, and control. The researchers developed nine stories of 250 words each. All stories had these components in this order: problem statement, irrelevant information, rule, distractors, less obvious critical information, and a question that called for text-based inference. The researchers pointed out that the contrived nature of the stories was necessary in order for them to look at the formal deduction type of inference that interested them. They noted that by creating these stories "much was sacrificed in terms of generalizability" (p. 336).

For both treatment groups, training occurred over 3 days with two of six training stories used each day. The transfer test (the remaining three stories) was given on the fourth day. In the systematic instruction group, students were first asked to read the story orally. If any words were misread by the student, the experimenter used language such as, "Stop, what word? The word is ____" (p. 337). For the first story, students were then asked "sequential questions . . . to elicit identification of the components of the story" (p. 326). Excerpts from one of these exchanges follows:

1. Statement of Topic

Experimenter: "Now let's go back to the beginning of the story so we can figure out the answer to the question at the end of the story. What does John want?"

Subject: "Not to be sick all the time."

2. Statement of Problem

Experimenter: "Right. What's important to him?"
Subject: "To eat good food with lots of vitamins."

3. Discrimination Questions

Experimenter: "Is getting big carrots most important to him?"
Subject: "No."
Experimenter: "Is getting sweet and tasty carrots most important to him?"
Subject: "No."
Experimenter: "Is getting cheap carrots most important to him?"
Subject: "No."

4. Review of Problem Statement

Experimenter: "What's most important to John?"
Subject: "To get food with lots of vitamins." (P. 337)

For the final questions, "the experimenter utilized a summary strategy, the purpose of which was to provide the subject practice in identifying the critical components and to focus the subject's attention on only those components" (p. 338).

For the story above, for example, the researcher asked what was most important to John, what foods would he chose, and why would he chose them.

For the second and third stories, the researcher "stopped the subject's reading at critical points in the story and asked the subject to identify the problem statement, rule and indirect information" (p. 338). For the fourth, fifth, and sixth stories, students were asked to read and answer the question. If the student got the answer correct, the researcher used the summary strategy "to verify the subject's ability to identify the various components of the story" (p. 338). If the student got the answer wrong, the experimenter used the questioning strategy used on the previous day.

In the corrective feedback group, students were asked to read the stories and answer the questions. The experimenter "corrected any misread words by providing a whole word correction" (p. 339). If the student got the question correct, the researcher asked why she or he made that choice. If the student "was able to cite verbally or point to the problem statement" (p. 339), the student was praised. If the student answered incorrectly, the researcher said, "No," and told the student the correct answer. Students in the control group did not receive any training.

On the third and last day of the study, all students read and answered a question about each of the remaining three stories. Carnine et al. (1982) found that the students in the systematic instruction group were better able to answer the end-of-story questions than were the students in the other treatment group or in the control group.

Rinehart, Stahl, and Erickson (1986) also focused on main idea training for sixth-grade students. In their study, they used a chapter from a sixth-grade social studies book. Seventy students of varied socioeconomic backgrounds participated in the study. Teachers received 90 minutes of training prior to the study and were given a script to use to provide summarization training to their students for 45–50 minutes a day for 5 consecutive days. Students in the control group read their basal and worked on worksheets, which was the regular instructional pattern for reading. The training components included explicit explanation of the skills, modeling, practice with feedback, and increasing task difficulty (short paragraphs first, later the chapter) and self-control training in which students took on increasing amounts of responsibility for task completion. The four skills taught were identifying and selecting main information, deleting trivial information, deleting redundant information, and relating the main idea and supporting information.

For both the pretests and the post-tests, students were asked to read a short article and write an outline for it. The outlining subtest of the Wisconsin Design for Reading Development (Otto, 1973) was used to assess the outline. There were no differences between treatment and control groups on this measure. For the post-tests, students were also asked to write summaries of short paragraphs. Findings revealed that the students in the treatment group were better able to identify main ideas than were students in the control group. There were no differences between the groups on their ability to identify minor information.

Across these studies, it is clear that fifth-grade and sixth-grade students given explicit instruction about how to identify main ideas do better on texts of main ideas than do students who are not provided with such instruction.

What Can Be Learned from the Comprehension Studies Included in the NRP Report?

Table 5 shows the conclusions that can reasonably be drawn across all the studies cited by the Comprehension subgroup (Kamil et al., 2000). What is clear across the studies is that researchers believe that children need to become strategic readers and that they need to learn from their teachers how to be strategic. What is not clear, from this body of research, is at what point such instruction is necessary. The studies in this report are predominately studies about students in third through sixth grade and make clear that students in those grade levels benefit from strategy instruction. There is, though, no information about how to help young children (K–2) make sense of texts. In a recent professional development endeavor with 120 literacy coaches and interventionists, we spent a day looking at data from 22 second-grade classrooms. Our interest was in understanding whether the students were paying attention to both meaning and phonics or predominantly to either meaning or phonics. We learned that most of the children in most of the classrooms paid more attention to phonics and considerably less to meaning to the degree that many of them were not understanding texts in which they were able to accurately read 90–94 percent of the words. This suggests to me that teachers need to gather information on the strategy use of all children (including those in K–2) and provide whole group, small group, and one-on-one instruction as needed. What is needed is research that documents the effects of doing so.

It is also important to note that for more than 30 years, as mentioned by the authors of the NRP report, researchers from various theoretical camps (see, for example, Clay, 1985; Durkin, 1978–1979; Goodman, Watson, & Burke, 1996) have been suggesting that during reading instruction teachers should provide students with the information they need to be strategic readers. Teachers, according to the NRP, have not heeded this call. The Comprehension subgroup (Kamil et al., 2000) suggests that perhaps teachers do not teach children how to read strategically because they have not been taught to do so. While this may or may not be true, it seems to be the responsibility of each and every teacher to learn how to teach children to read (rather than, as is common, to listen and then ask questions to assess comprehension). The body of research cited by Kamil et al. provides some information from experimental and quasi-experimental research about how to do this with students in grades 3 to 6. K–2 teachers might consider action research (see, for example, Taberski, 2000) as a source of information for how they might proceed.

Table 5. Instructional Implications from Comprehension Studies as Cited in the NRP Report

If you want students in grades 3 to 6 to . . .	*You could . . .*
Monitor comprehension	Teach them to monitor (see, for example, Baumann et al., 1992; Block, 1993; Tregaskes & Daines, 1989; Paris et al., 1984)

Know and use a variety of reading strategies, learn more from what they read, and do better on standardized tests	Provide them with teachers who focus instruction on reading strategies across the year (Brown et al., 1996)
Better remember what they read	Give them graphic organizers (e.g., Alvermann & Boothby, 1986; Darch et al., 1986)
Do better on standardized tests	Have them engage in Reciprocal Teaching in which students are scaffolded into learning how to clarify, question, summarize, and predict (Lysynchuk et al., 1990; Palinscar & Brown, 1984)
Better answer questions about a passage	Teach them ways to find answers (e.g., Raphael & Pearson, 1985; Raphael & McKinney, 1983; Raphael & Wonnacott, 1985; Seretny & Dean, 1986; Garner et al., 1984)
Do better on answering inferential questions	Teach them how to ask inferential questions (Davey & McBride, 1986); teach them cloze procedures, a checklist, or structured overviews (Carr et al., 1991); teach them a reciprocal inference procedure (Reutzel & Hollingsworth, 1991a)
Do better on answering questions about a story	Teach them how to use their knowledge of story structure to ask their own questions (Nolte & Singer, 1985; Short & Ryan, 1984)
Do better on identifying main ideas	Provide direct instruction followed by collaborative practice (Stevens et al., 1987, 1991)
	Provide direct instruction in identifying the main idea (e.g., Rinehart et al., 1986; Carnine et al., 1982)

Helpful Quotes about Comprehension Instruction from the NRP Report
As teachers work out for themselves (and address with others) the instructional implications of the comprehension studies cited by the NRP, these comments from the report might prove useful:

> Readers normally acquire strategies for active comprehension informally. (P. 4-40)
>
> Comprehension strategies are specific procedures that guide students to become aware of how well they are comprehending as they attempt to read and write. (P. 4-40)

When faced with a comprehension problem, a good strategy user will coordinate strategies and shift strategies as it is appropriate to do so. They will constantly alter, adjust, modify, and test until they construct meaning and the problem is solved. (P. 4-47)

The general finding is that when readers are given cognitive strategy instruction, they make significant gains on measures of reading comprehension over students trained with conventional instruction procedures. (P. 4-40)

The data suggests that students at all skill levels would benefit from being taught . . . strategies (Rosenshine, Meister, & Chapman, 1996, p. 201). (P. 4-46)

[E]ffective reading instruction is associated more with independent teacher action than with implementation of basal text prescriptions. (P. 4-48)

[D]eveloping metacognitive readers who understand their reasoning requires teachers who themselves understand their reasoning, as well as a supportive environment in the schools for strategy learning. (P. 4-48)

[S]trategies are not skills that can be taught by drill; they are plans for constructing meaning. Teaching students to acquire and use strategies may require altering traditional approaches to strategy instruction. It may be necessary to free teachers of the expectation that their job is to follow directions narrowly. (P. 4-47)

Strategic reading requires strategic teaching, which involves putting teachers in positions where their minds are the most valued educational resource. (P. 4-49)

7: How Should Teachers Use the Findings?

In order to determine the usefulness of the NRP findings for their classrooms, teachers need to interrogate the findings in at least five ways.

First, whenever possible, teachers need to look at the data by grade level (e.g., preschool versus kindergarten) and by population characteristics (e.g., at-risk versus normally achieving students). Keep in mind that a meta-analysis, the sole tool for the PA and phonics study, is designed to reach broad conclusions across diverse studies. However, the findings for all the studies included may not match the findings for a particular age group or for particular types of children. For example, there are 5 studies of PA training involving preschoolers. Three of the studies concluded that children in treatment groups do better on tests of PA that measure what they explicitly were taught than do children who were not taught those skills (Byrne & Fielding-Barnsley, 1991, 1993, 1995; Haddock, 1976; Treiman & Baron, 1983). In the fourth study by Fox and Routh (1976), the experimenters gave 20 preschool children "Gibson letter-like symbols" (p. 71) and explained to them that the symbols "said" sounds of letters. The children were then taught to blend the symbols to make words. For example, the children were told that one symbol "said" /m/ as in *may* while another one "said" /a/ as in *may* and were taught to blend the two made-up symbols to "say" /may/. The researchers chose symbols rather than letters so that the outcome would not be affected by differences in what the children knew about letters and sounds. This study concluded that blending training helps children who already know how to segment. The fifth study, Solity (1996), taught two groups of preschoolers phonological skills. In one group, the reading group, the children were taught skills "exclusively in the context of print" and engaged in activities such as reading sentences and short stories, talking about the stories, rhyming by changing the initial phoneme, and blending sounds to make whole words (p. 108). In the second group, the phonology group, children engaged in the same activities, but "their work involved manipulating sounds only and was not related to print" (p .108). Solity found that the children in the reading group did better than the children in the phonology group on tasks that involved reading sight words and letter-sound recognition.

In the meta-analysis, the findings from these 5 studies with preschoolers were combined with the studies of children at other grade levels, with studies of at-risk children, and with studies of children who are learning a language other than English in a country other than the United States, and from this, broad generalizations were made. For example, one generalization about PA studies is that it is better to teach blending and segmenting together rather than teach either alone. These generalizations are patterns found across studies and not patterns that apply to all children involved in the studies. Teachers need to try to ascertain what the NRP found about each of the elements (PA, phonics, fluency, vocabulary, and comprehension) at each grade level in order to determine if the patterns hold for their grade level and also hold for the children with whom they work. Generalizations may work well for researchers, but teachers need to know what the research results are relative to the children they teach.

Second, teachers need to think carefully about the meaning of the statistical significance reported by the researchers. In the Fox and Routh (1976) study, for example, children who were "proficient at segmenting" (p. 70) before the treatment had

a gain effect of 1.61, suggesting that the treatment helped them with the assessment measure. However, children who were not previously considered proficient at segmenting were not helped by the treatment. Indeed, the gain effect for that group was −.1. When these two gain effects are combined together to get a gain effect of 1.51, it is easy to miss the fact that the treatment worked only for children who already knew how to segment and that it had a negative influence on the children who did not previously know how to segment.

Third, teachers need to gather information about the assessment tools and consider the language used in summarizing findings from the assessments. In the Fox and Routh (1976) study mentioned above, the children first learned to associate the Gibson letter-like symbols with sounds. There were five symbols; the five sounds were "*a* as in *may*, *m* as in *may*, *e* as in *me*, *s* as in *see* and *w* as in *we*" (p. 72). Next,

> the five letter-like forms were combined to form two lists of three words each. The words for list A were *me*, *see*, and *way*, and the words for list B were *we*, *say*, and *may*. (P. 71)

The children who knew how to segment before the treatment began (the "segmenters") did better on this test than did the children who did not know how to segment when the treatment began. They also did better than the children who did not receive a treatment. This task was considered by the NRP Alphabetics subgroup to be a *reading* task, and so this study contributed to their finding that PA training helps with *reading.* In interrogating the findings of the Fox and Routh (1976) study, teachers need to consider whether they believe that asking children to read words made up of the same Gibson letter-like symbols they had just learned would be a *reading* assessment. Similar questions need to be raised for all assessments in all studies.

Fourth, teachers should also consider the practical significance of the findings for the classroom. In considering practical significance, teachers need to look at the amount of time that the intervention took and consider that investment of time relative to the gain. Consider, for example, the Ball and Blachman (1991) study of PA training. In this study, 89 kindergartners were placed in one of three groups—phoneme segmentation plus letter-name and letter-sound training; letter-name and letter-sound training; and no training. The children in the training groups received instruction 20 minutes a day, 4 days a week, for 7 weeks. The children who were in the phoneme segmentation group did better on phoneme segmentation tasks than did the other two groups. All three groups scored similarly on letter names. The children in the two treatment groups did better on letter sounds than did the control group. The phoneme segmentation group did "better" on the Woodcock-Johnson. Here, "better" means that 37 percent of the children in the phoneme segmentation group could read 4 words on the Woodcock-Johnson while only 13 percent of the nonphoneme segmentation treatment group and only 7 percent of the control group could. Teachers need to weigh for themselves whether spending 9+ hours spread over 7 weeks balances the outcome of having kindergarten children able to read 4 words on the Woodcock-Johnson test and to score higher on tests matched to what was taught.

Fifth, teachers need to consider the practical significance of the treatment itself. I do not imagine that many teachers would consider the Fox and Routh (1976) study—

which taught children to blend made-up symbols—to have practical applications for the preschool classroom. Indeed, teachers might worry that such instruction would confuse children when they subsequently were taught that *m* "said" /m/. Solity's study for preschoolers, on the other hand, does seem to have practical applications for the preschool classroom as it suggests that children learn PA best in when it is taught as part of reading.

Sixth, teachers need to think about how their beliefs align with the beliefs of the members of the various subgroups. Take, for example, differences of opinion about training in PA. Research has shown that there is a correlation between high PA scores and later reading achievement. However, just because two things are correlated does not mean that one causes the other. My favorite way of explaining this comes from Coles (2003). He notes that children who live in particular zip codes often have higher standardized test scores than do children who live in other zip codes. However, test scores would not be improved by changing the zip codes of the lower-achieving children.

This applies to PA in the following way: some teachers and researchers believe that since there is a correlation between PA and later reading achievement, teachers should provide explicit instruction in PA and that instruction will lead to improved reading achievement. These individuals would value the PA studies showing that PA skills can be taught, that is, that children receiving PA training do better on tests of PA than do children who have not had this training.

Other teachers and researchers believe that other factors explain the correlational relationship (see, for example, Krashen, 2004). One theory is that children who have been read to a lot and have played language games at home have higher PA scores *and* end up scoring higher on reading comprehension measures. These individuals who hold this belief would not value highly the outcomes of PA training but would be more interested in studies that looked at how being read to a lot and playing language games helped both with PA and, later, with reading comprehension. Teachers need to consider their beliefs relative to these kinds of issues and use that reflection to inform how they value the findings from the studies included in the NRP report.

8: Summary and So What?

Reading the NRP report is a daunting task—449 pages of a not-so-well-edited, sometimes contradictory narrative about 441 studies related to reading.

Reading most of the articles cited in the NRP report to write this "short" book has also been a daunting task.

Reading this publication is perhaps not as daunting, but there is an enormous amount of detail to sort through—and most of it is just not that fascinating (or pleasant) to read through. It is a dense reference book.

So where does that leave teachers?

First, I would suggest that teachers obtain their free copy of the NRP report and skim it to understand the organizational structure (a necessity as the pages are not numbered consecutively but by section) and be able to find the pages referenced in this and other reports on it.

Second, teachers should read the Ever-So-Brief Summary below as well as the Executive Summary found at the beginning of this teacher's guide. Also, take a look at the summary charts about each aspect of reading discussed in the report and at the quotes from the NRP report included at the end of each chapter. These "quick bites" will provide an overview of a very extensive body of literature.

Third, teachers should read further about classroom-based implications of these findings and also read books and articles written by teachers that show what practice looks like when the findings from these (and other) studies are incorporated into K–12 classrooms. A good place to start would be to read some of the books written about best practices: for K–3, *On Solid Ground: Strategies for Teaching Reading K–3* (Taberksi, 2000); for upper elementary, *Still Learning to Read: Teaching Students in Grades 3–6* (Sibberson & Szymusiak, 2003); for middle school and high school, Beers's (2002) *When Kids Can't Read: What Teachers Can Do: A Guide for Teachers 6–12*; and Tovani's (2000) *I Read It, but I Don't Get It: Comprehension Strategies for Adolescent Readers*.

Fourth, teachers should read beyond the NRP report. The field of reading understands a great deal about the reading process. This understanding is research-based. It is broader than the research reported in the NRP report. It includes research well suited to educational purposes—case studies, classroom-based research, correlational studies, and so on. A good place to begin would be with Allington's *What Really Matters for Struggling Readers: Designing Research-Based Programs* (2005b).

Fifth, teachers should use their knowledge base—of reading and of each and every child in their classroom—to make informed decisions about the best practices for their classrooms. For far too long, individuals outside the classroom have attempted to control, via mandates and textbook adoptions, what it is that teachers do in the name of reading instruction. Teachers know their students and need to know in depth about the reading process, reading strategies, reading instruction, and children's literature. If teachers would commit to broadening and deepening their knowledge base and using that knowledge and their knowledge of particular children to inform instruction, we could change educational outcomes across the country.

Sixth, there is only one study included in the phonics, PA, fluency, comprehension, and vocabulary section of the NRP report that extensively details what

the classroom teachers of the students in these studies knew in depth and put into practice and how that impacted their students. That is the study by Brown et al. (1996) about SAIL teachers. I would encourage teachers across the country to start naming what it is they know and believe, to look at the relationship between those beliefs and their practices, and to collect information about student progress across the year. It is not enough to argue that we want control over classroom instruction; we need to document how knowledgeable teachers impact student achievement. We change lives every day, every year. We need to show others that we do so.

Seventh, teachers should share what they know with other teachers so that together teachers can broaden and deepen their knowledge base, strengthen their practices, and form a community of professionals. A good place to start would be to join a professional organization such as the National Council of Teachers of English.

Eighth, teachers need to get their voices heard in the conversations about best practices at the local, state, and national levels. This is particularly important at this writing, October 2007, as the federal government takes steps to reauthorize the Elementary and Secondary Education Act, better known as No Child Left Behind. In order to be heard, teachers need to be talking and writing and faxing and emailing. And their voices need to be well-informed voices. The education of K–12 children is at stake, and those of us who spend our lives working with those children are the ones whose voices should be heard loudest and whose voices should be shaping public policy. It is time to reverse the pattern of having those outside K–12 education determine what it is that teachers ought to be doing. It is a time for educators to seize both responsibility and accountability for education.

Across all of these suggestions, keep in mind that it is the teacher who makes the difference. We each need to commit to making that difference. We can do this by expanding our knowledge base and asserting our role as curricular experts. Teachers can help each and every child read. We just need to do this and let others know we can.

An Ever-So-Brief Summary with Book Recommendations

1. **Phonemic Awareness.** According to the studies cited in the NRP report, this is best taught to very young children (K–1) using letters, and when letters are used, PA instruction is considered to be phonics. Therefore, it is not necessary to have a separate instructional time for PA. Rather, children should have opportunities to learn about how language is made up of parts (e.g., onsets and rimes, or word families) as part of phonics instruction. An effective way to do this in the classroom? Provide time for students to write using invented spelling (pp. 2-1 through 2-86). (See Strickland, 1998, for further information about invented spelling.)

2. **Phonics.** According to the studies cited in the NRP report, there is no evidence that phonics instruction helps in kindergarten or in grades 2 to 9. It does help first graders learn the alphabetic principle—that there is a relationship between letters and sounds. No one method is better than any other. For example, for at-risk first graders, a modified whole language approach and one-on-one Reading Recovery–like instruction both helped children with comprehension (pp. 2-89 through 2-176). This phonics instruction should be conducted in the context of whole, meaningful text. (See Moustafa, 1997, for information on embedded, whole-part-whole instruction.)

3. **Fluency.** According to the authors of the Fluency report, the practice of round robin (at any age) does not help children and can indeed hurt them. However, according to the studies cited in the Fluency report, repeated oral reading (K–12) helps with comprehension because reading to readers fluidly instead of word-by-word reading helps them better understand the text. Ways to help with this? Try such things as readers theater (pp. 3-1 through 3-43). (See Opitz and Rasinski's *Good-bye Round Robin: 25 Effective Oral Reading Strategies* [1998] for additional instructional suggestions.)

4. **Vocabulary (grades 3 to 8).** One method is not better than another. Children learn most of their vocabulary incidentally (pp. 4-15 through 4-35). (For further information about vocabulary learning, see Nagy, 1988.)

5. **Comprehension (grades 3 to 6).** Children need to learn that print makes sense and to develop a variety of strategies for making sense of print (pp. 4-39 through 4-168). (For further information on teaching for comprehension, see the references listed in Chapter 8: Beers, 2002; Sibberson & Szymusiak, 2003; Taberski, 2000; Tovani, 2000; see also Harvey & Goudvis, 2000.)

Across all of these recommendations? According to the studies cited in the NRP report, if we want children to learn something, we need to teach them that something. Want great readers? Then teach children what great readers do.

References

Adams, A., Carnine, D., & Gersten, R. (1982). Instructional strategies for studying content area texts in the intermediate grades. *Reading Research Quarterly, 18*, 27–55.

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

Allington, R. L. (1977). If they don't read much, how they ever gonna get good? *Journal of Reading, 21*, 57–61.

Allington, R. L. (Ed.). (2002). *Big brother and the National Reading Curriculum: How ideology trumped evidence.* Portsmouth, NH: Heinemann.

Allington, R. L. (2005a). The other five "pillars" of effective reading instruction. *Reading Today, 22*(6), 3, 5–6.

Allington, R. L. (2005b). *What really matters for struggling readers: Designing research-based programs* (2nd ed.). Boston: Allyn and Bacon.

Alvermann, D. E., & Boothby, P. R. (1986). Children's transfer of graphic organizer instruction. *Reading Psychology, 7*, 87–100.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the Commission on Reading.* Washington, DC: National Academy of Education and National Institute of Education; Champaign, IL: Center for the Study of Reading.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 255–291). New York: Longman.

Anderson, R. C., Wilkinson, I. G., & Mason, J. M. (1991). A microanalysis of the small-group, guided reading lesson: Effects of an emphasis on global story meaning. *Reading Research Quarterly, 26,* 417–441.

Anderson, V., & Roit, M. (1993). Planning and implementing collaborative strategy instruction for delayed readers in grades 6–10. *Elementary School Journal, 94*, 121–137.

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26*, 49–66.

Barker, T. A., & Torgesen, J. K. (1995). An evaluation of computer-assisted instruction in phonological awareness with below average readers. *Journal of Educational Computing Research, 13*, 89–103.

Baumann, J. F., Seifer-Kessell, N., & Jones, L. A. (1992). Effect of think-aloud instruction on elementary students' comprehension monitoring abilities. *Journal of Reading Behavior, 24*, 143–172.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). The effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*, 506–521.

Beers, K. (2002). *When kids can't read: What teachers can do: A guide for teachers 6–12.* Portsmouth, NH: Heinemann.

Blachman, B. A., Ball, E. W., Black, R. S., & Tangel, D. M. (1994). Kindergarten teachers develop phoneme awareness in low-income, inner-city classrooms:

Does it make a difference? *Reading and Writing: An Interdisciplinary Journal, 6*, 1–18.

Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R., & McGraw, C. K. (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*, 239–273.

Block, C. C. (1993). Strategy instruction in a literature-based reading program. *Elementary School Journal, 94*, 139–151.

Blum, I. H., Koskinen, P. S., Tennant, N., Parker, E. M., Straub, M., & Curry, C. (1995). Using audiotaped books to extend classroom literacy instruction into the homes of second-language learners. *Journal of Reading Behavior, 27*, 535–563.

Bond, C. L., Ross, S. M., Smith, L. J., & Nunnery, J. A. (1995–1996). The effects of the sing, spell, read and write program on reading achievement of beginning readers. *Reading Research and Instruction, 35*, 122–141.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301*, 419–421.

Bradley, L., & Bryant, P. E. (1985). *Rhyme and reason in reading and spelling* (International Academy for Research in Learning Disabilities Monograph Series No. 1). Ann Arbor: University of Michigan Press.

Brady, S., Fowler, A., Stone, B., & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia, 44*, 26–59.

Bramlett, R. K. (1994). Implementing cooperative learning: A field study evaluating issues for school-based consultants. *Journal of School Psychology, 32*, 67–84.

Brennan, F., & Ireson, J. (1997). Training phonological awareness: A study to evaluate the effects of a program of metalinguistic games in kindergarten. *Reading and Writing: An Interdisciplinary Journal*, *9*, 241–263.

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

Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology, 88*, 18–37.

Byrne, B., & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology*, *83*, 451–455.

Byrne, B., & Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A 1-year follow-up. *Journal of Educational Psychology, 85,* 104–111.

Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3- year follow-up and a new preschool trial. *Journal of Educational Psychology, 87,* 488–503.

Camilli, G., Vargas, S., & Yurecko, M. (2003, May 8). Teaching children to read: The fragile link between science and federal education policy. *Education Policy Analysis Archives, 11*(15). Retrieved May 9, 2003, from http://epaa.asu.edu/epaa/v11n15/

Canney, G., & Winograd, P. (1979). *Schemata for reading and reading comprehension performance* (Technical Report No. 120). Urbana, IL: Center for the Study of Reading. (ERIC Document Reproduction Service No. ED169520)

Carnine, D. W., Kameenui, E. J., & Woolfson, N. (1982). Training of textual dimensions related to text-based inferences. *Journal of Reading Behavior*, *14*, 335–340.

Carr, E., Bigler, M., & Morningstar, C. (1991). The effects of the CVS strategy on children's learning. In J. Zutell & S. McCormick (Eds.), *Learner factors/teacher factors: Issues in literacy research and instruction* (Fortieth Yearbook of the National Reading Conference, pp .193–200). Chicago: National Reading Conference.

Carver, R. P., & Hoffman, J. V. (1981). The effect of practice through repeated reading on gain in reading ability using a computer-based instructional system. *Reading Research Quarterly*, *16*, 374–390.

Castle, J. M., Riach, J., & Nicholson, T. (1994). Getting off to a better start in reading and spelling: The effects of phonemic awareness instruction within a whole language program. *Journal of Educational Psychology*, *86*, 350–359.

Clay, M. M. (1985). *The early detection of reading difficulties* (3rd ed.). Portsmouth, NH: Heinemann.

Coles, G. (2003). *Reading the naked truth: Literacy, legislation and lies.* Portsmouth, NH: Heinemann.

Cunningham, A. E. (1990). Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, *50*, 429–444.

Darch, C. B., Carnine, D. W., & Kameenui, E. J. (1986). The role of graphic organizers and social structure in content area instruction. *Journal of Reading Behavior*, *18*, 275–295.

Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. *Journal of Educational Psychology, 78*, 256–262.

Davidson, M., & Jenkins, J. R. (1994). Effects of phonemic processes on word reading and spelling. *Journal of Educational Research*, *87*, 148–157.

Dickinson, D. K., & Smith, M. W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly*, *29*, 104–122.

Dixon-Krauss, L. A. (1995). Partner reading and writing: Peer social dialogue and the zone of proximal development. *Journal of Reading Behavior*, *27*, 45–63.

Dowhower, S. L. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, *22*, 389–406.

Durkin, D. (1978–1979). What classroom observations reveal about reading comprehension. *Reading Research Quarterly*, *14*, 481–533.

Durkin, D. (1993). *Teaching them to read* (6th ed.). Boston: Allyn and Bacon.

Ehri, L., Correro, G., Shanahan, T., Willows, D., & Yatvin, J. (2000). Alphabetics. In National Institute of Child Health and Human Development, *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754, pp. 2-1–2-176). Washington, DC: U.S. Government Printing Office.

Ehri, L. C., & Wilce, L. S. (1987). Does learning to spell help beginners learn to read words? *Reading Research Quarterly*, *22*, 47–65.

Eldredge, J. L. (1990). Increasing the performance of poor readers in the third grade with a group-assisted strategy. *Journal of Educational Research*, *84*, 69–77.

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

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

Farmer, A. R., Nixon, M., & White, R. T. (1976). Sound blending and learning to read: An experimental investigation. *British Journal of Educational Psychology*, *46*, 155–163.

Foorman, B. R., Francis, D. J., Fletcher, J. M., 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.

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

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

Fox, B., & Routh, D. K. (1976). Phonemic analysis and synthesis as word-attack skills. *Journal of Educational Psychology*, *68*, 70–74.

Fox, B., & Routh, D. K. (1984). Phonemic analysis and synthesis as word attack skills: Revisited. *Journal of Educational Psychology*, *76*, 1059–1064.

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

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

Garan, E. M. (2002). *Resisting reading mandates: How to triumph with the truth.* Portsmouth, NH: Heinemann.

Garner, R., Chou Hare, V., Alexander, P., Haynes, J., & Winograd, P. (1984). Inducing use of a text lookback strategy among unsuccessful readers. *American Educational Research Journal*, *21*, 789–798.

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

Gilbert, L. M., Williams, R. L., & McLaughlin, T. F. (1996). Use of assisted reading to increase correct reading rates and decrease error rates of students with learning disabilities. *Journal of Applied Behavior Analysis*, *29*, 255–257.

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.

Goodman, Y. M., & Burke, C. L. (1972). *Reading miscue inventory manual: Procedure for diagnosis and evaluation.* New York: Macmillan.

Goodman, Y. M., Watson, D. J., & Burke, C. L. (1996). *Reading Strategies: Focus on Comprehension*. Katonah, NY: Richard C. Owen.

Greaney, K. T., Tunmer, W. E., & Chapman, J. W. (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.

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

Gross, J., & Garnett, J. (1994). Preventing reading difficulties: Rhyme and alliteration in the real world. *Educational Psychology in Practice*, *9*, 235–240.

Guthrie, J. T., Van Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C. C., et al. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly*, *31*, 306–332.

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

Harris, T. L., & Hodges, R. E. (Eds.). (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.

Harste, J. C. (1985). Portrait of a new paradigm: Reading comprehension research. In A. Crismore (Ed.), *Landscapes: A state-of-the-art assessment of reading comprehension research, 1974–1984 final report* (pp. 1–24). USDE-C-300-83-0130. Bloomington, IN: Indiana University.

Harvey, S., & Goudvis, A. (2000). *Strategies that work: Teaching comprehension to enhance understanding.* Portland, ME: Stenhouse.

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

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, *65*, 41–57.

Hohn, W. E., & Ehri, L. C. (1983). Do alphabet letters help prereaders acquire phonemic segmentation skill? *Journal of Educational Psychology, 75*, 752–762.

Hollingsworth, P. M. (1978). An experimental approach to the impress method of teaching reading. *Reading Teacher*, *31*, 624–626.

Hurford, D. P., Johnston, M., Nepote, P., Hampton, S., Moore, S., Neal, J., et al. (1994). Early identification and remediation of phonological-processing deficits in first-grade children at risk for reading disabilities. *Journal of Learning Disabilities*, *27*, 647–659.

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

Judy, J. E., Alexander, P. A., Kulikowich, J. M., & Wilson, V. L. (1988). Effects of two instructional approaches and peer tutoring on gifted and nongifted sixth-grade students' analogy performance. *Reading Research Quarterly*, *23*, 236–256.

Kamil, M. L., Ferguson, G., Garza, N., Trabasso, T., & Williams, J. (2000). Comprehension. In National Institute of Child Health and Human Development, *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its*

implications for reading instruction: Reports of the subgroups (NIH Publication No. 00-4754, pp. 4-1–4-168). Washington, DC: U.S. Government Printing Office.

Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, J. (1994). Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis*, *27*, 49–61.

Keats, E. J. (1976). *The Snowy Day.* New York: Puffin Books.

Kennedy, K. M., & Backman, J. (1993). Effectiveness of the Lindamood Auditory Discrimination in Depth Program with students with learning disabilities. *Learning Disabilities Research and Practice, 8*, 253–259.

Klesius, J. P., Griffith, P. L., & 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.

Klingner, J. K., Vaughn, S., & Schumm, J. S. (1998). Collaborative strategic reading during social studies in heterogeneous fourth-grade classrooms. *Elementary School Journal*, *99*, 3–22.

Krashen, S. (2001). Does “pure” phonemic awareness training affect reading comprehension? *Perceptual and Motor Skills*, *93*, 356–358.

Krashen, S. (2004). False claims about literacy development. *Educational Leadership*, *61*(6), 18–21.

Leach, D. J., & Siddall, S. W. (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.

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

Liddle, W. (Ed.). (1977). *Reading for concepts: Book E* (2nd ed.). New York: McGraw-Hill.

Lorenz, L., & Vockell, E. (1979). Using the neurological impress method with learning disabled readers. *Journal of Learning Disabilities*, *12*, 420–422.

Lovett, M. W., Barron, R. W., Forbes, J. E., Cuksts, B., & Steinbach, K. A. (1994). Computer speech-based training of literacy skills in neurologically impaired children: A controlled evaluation. *Brain and Language*, *47*, 117–154.

Lovett, M. W., Lacerenza, L., Borden, S. L., Frijters, J. C., Steinbach, K. A., & De Palma, M. (2000). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. *Journal of Educational Psychology*, *92*, 263–283.

Lovett, M. W., Ransby, M. J., Hardwick, N., Johns, M. S., & Donaldson, S. A. (1989). Can dyslexia be treated? Treatment-specific and generalized treatment effects on dyslexic children’s response to remediation. *Brain and Language*, *37*, 90–121.

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

Lum, T., & Morton, L. L. (1984). Direct instruction in spelling increases gain in spelling and reading skills. *Special Education in Canada*, *58*(2), 41–45.

Lysynchuk, L. M., Pressley, M., & Vye, N. J. (1990). Reciprocal teaching improves standardized reading-comprehension performance in poor comprehenders. *Elementary School Journal*, *90*, 469–484.

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.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, *48*, 986–992.

Markman, E. M. (1981). Comprehension monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills* (pp. 61–84). New York: Academic Press.

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

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

McGuinness, D., McGuinness, C., & Donohue, J. (1995). Phonological training and the alphabet principle: Evidence for reciprocal causality. *Reading Research Quarterly*, *30*, 830–852.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly*, *20*, 522–535.

Moustafa, M. (1997). *Beyond traditional phonics: Research discoveries and reading instruction.* Portsmouth, NH: Heinemann.

Murray, B. A. (1998). Gaining alphabetic insight: Is phoneme manipulation skill or identity knowledge causal? *Journal of Educational Psychology*, *90*, 461–475.

Nagy, W. E. (1988). *Teaching vocabulary to improve reading comprehension*. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills and National Council of Teachers of English; Newark, DE: International Reading Association.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*, 237–270.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

Neill, K. (1980). Turn kids on with repeated readings. *Teaching Exceptional Children*, *12*, 63–64.

Nolte, R. Y., & Singer, H. (1985). Active comprehension: Teaching a process of reading comprehension and its effects on reading achievement. *Reading Teacher*, *39*, 24–31.

O'Connor, R. E., & Jenkins, J. R. (1995). Improving the generalization of sound/symbol knowledge: Teaching spelling to kindergarten children with disabilities. *Journal of Special Education*, *29*, 255–275.

O'Connor, R. E., Jenkins, J. R., & Slocum, T. A. (1995). Transfer among phonological tasks in kindergarten: Essential instructional content. *Journal of Educational Psychology, 87*, 202–217.

O'Connor, R. E., Notari-Syverson, A., & Vadasy, P. F. (1996). Ladders to literacy: The effects of teacher-led phonological activities for kindergarten children with and without disabilities. *Exceptional Children, 63*, 117–130.

O'Connor, R. E., Notari-Syverson, A., & Vadasy, P. F. (1998). First-grade effects of teacher-led phonological activities in kindergarten for children with mild disabilities: A follow-up study. *Learning Disabilities Research and Practice*, *13*, 43–52.

Opitz, M. F., & Rasinski, T. V. (1998). *Good-bye round robin: 25 effective oral reading strategies.* Portsmouth, NH: Heinemann.

Otto, W. (1973). *Wisconsin tests of reading skill development: Study skills.* Minneapolis, MN: National Computer Systems.

Palinscar, A. S., & Brown, A. I. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*, 117–175.

Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*, 1239–1252.

Pearson, P. D., & Johnson, D. D. (1978). *Teaching reading comprehension.* New York: Holt, Rinehart & Winston.

Raphael, T. E., & McKinney, J. (1983). An examination of fifth- and eighth-grade children's question-answering behavior: An instructional study in metacognition. *Journal of Reading Behavior*, *15*, 67–86.

Raphael, T. E., & Pearson, P. D. (1985). Increasing students' awareness of sources of information for answering questions. *American Educational Research Journal*, *22*, 217–235.

Raphael, T. E., & Wonnacott, C. A. (1985). Heightening fourth-grade students' sensitivity to sources of information for answering comprehension questions. *Reading Research Quarterly*, *20*, 282–296.

Rasinski, T. V. (1990). Effects of repeated reading and listening-while-reading on reading fluency. *Journal of Educational Research*, *83*, 147–150.

Rasinski, T. V., Padak, N., Linek, W., & Sturtevant, E. (1994). Effects of fluency development on urban second-grade readers. *Journal of Educational Research*, *87*, 158–165.

Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-while-listening, and independent reading with computer-based speech feedback. *Reading Research Quarterly*, *23*, 219–235.

Reitsma, P., & Wesseling, R. (1998). Effects of computer-assisted training of blending skills in kindergartners. *Scientific Studies of Reading*, *2*, 301–320.

Reutzel, D. R., & Hollingsworth, P. M. (1991a). Reading comprehension skills: Testing the distinctiveness hypothesis. *Reading Research and Instruction*, *30*, 32–46.

Reutzel, D. R., & Hollingsworth, P. M. (1991b). Reading time in school: Effect on fourth graders' performance on a criterion-referenced comprehension test. *Journal of Educational Research*, *84*, 170–176.

Rinehart, S. D., Stahl, S. A., & Erickson, L. G. (1986). Some effects of summarization training on reading and studying. *Reading Research Quarterly, 21*, 422–438.

Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology*, *86*, 54–64.

Rose, T. L., & Beattie, J. R. (1986). Relative effects of teacher-directed and tape previewing on oral reading. *Learning Disability Quarterly*, *9*, 193–199.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*, 479–530.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66*, 181–221.

Samuels, S. J., Shanahan, T., & Shaywitz, S. (2000). Fluency. In National Institute of Child Health and Human Development, *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754, pp. 3-1–3-43). Washington, DC: U.S. Government Printing Office.

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

Schreiber, P.A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior*, *12*, 177–186.

Schreiber, P. A. (1987). Prosody and structure in children's syntactic processing. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 243–270). New York: Academic Press.

Senechal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, *28*, 360–374.

Seretny, M. L., & Dean, R. S. (1986). Interspersed post passage questions and reading comprehension achievement. *Journal of Educational Psychology*, *78*, 228–229.

Shany, M. T., & Biemiller, A. (1995). Assisted reading practice: Effects on performance for poor readers in grades 3 and 4. *Reading Research Quarterly*, *30*, 382–395.

Short, E. J., & Ryan, E. B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. *Journal of Educational Psychology*, *76*, 225–235.

Sibberson, F., & Szymusiak, K. (2003). *Still Learning to Read: Teaching Students in Grades 3–6*. Portland, ME: Stenhouse.

Silberberg, N. E., Iversen, I. A., & Goins, J. T. (1973). Which remedial reading method works best? *Journal of Learning Disabilities*, *6*, 547–556.

Simmons, D. C., Fuchs, L. S., Fuchs, D., Mathes, P., & Hodge, J. P. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning-disabled and low-performing students in regular classrooms. *Elementary School Journal*, *95*, 387–408.

Sindelar, P. T., Monda, L. E., & O'Shea, L. J. (1990). Effects of repeated readings on instructional- and mastery-level readers. *Journal of Educational Research*, *83*, 220–226.

Smith, D. D. (1979). The improvement of children's oral reading through the use of teacher modeling. *Journal of Learning Disabilities*, *12*, 172–175.

Smith, K., Johnson, D. W., & Johnson, R. T. (1981). Can conflict be constructive? Controversy versus concurrence seeking in learning groups. *Journal of Educational Psychology*, *73*, 651–663.

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

Solity, J. (1996). Phonological awareness: Learning disabilities revisited? *Educational and Child Psychology*, *13*, 103–113.

Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher*, *9*(11), 11–16.

Stevens, R. J. (1988). Effects of strategy training on the identification of the main idea of expository passages. *Journal of Educational Psychology*, *80*, 21–26.

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative integrated reading and composition: Two field experiments. *Reading Research Quarterly*, *22*, 433–454.

Stevens, R. J., Slavin, R. E., & Farnish, A. M. (1991). The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology*, *83*, 8–16.

Strickland, D. S. (1998). *Teaching phonics today: A primer for educators*. Newark, DE: International Reading Association.

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

Taberksi, S. (2000). *On solid ground: Strategies for teaching reading K–3*. Portsmouth, NH: Heinemann.

Tangel, D. M., & Blachman, B. A. (1992). Effect of phoneme awareness instruction on kindergarten children's invented spelling. *Journal of Reading Behavior*, *24*, 233–261.

Taylor, B. M., & Frye, B. J. (1992). Comprehension strategy instruction in the intermediate grades. *Reading Research and Instruction, 32*, 39–48.

Torgesen, J. K., Morgan, S. T., & Davis, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology, 84*, 364–370.

Torgesen, J. K., Wagner, R. K, Rashotte, C. A., 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.

Tovani, C. (2000). *I read it, but I don't get it: Comprehension strategies for adolescent readers*. Portland, ME: Stenhouse.

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

Tregaskes, M. R., & Daines, D. (1989). Effects of metacognitive strategies on reading comprehension. *Reading Research and Instruction*, *29*, 52–60.

Treiman, R., & Baron, J. (1983). Phonemic-analysis training helps children benefit from spelling-sound rules. *Memory and Cognition*, *11*, 382–389.

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

Uhry, J. K., & Shepherd, M. J. (1993). Segmentation/spelling instruction as part of a first-grade reading program: Effects on several measures of reading. *Reading Research Quarterly*, *28*, 219–233.

Umbach, B. T., Darch, C. B., & Halpin, G. (1989). Teaching reading to low performing first graders: A comparison of two instructional approaches. *Journal of Instructional Psychology*, *16*, 112–21.

Vadasy, P. F., Jenkins, J. R., Antil, L. R., Wayne, S. K., & O'Connor, R. E. (1997). Community-based early reading intervention for at-risk first graders. *Learning Disabilities Research and Practice, 12*, 29–39.

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

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

Warrick, N., Rubin, H., & Rowe-Walsh, S. (1993). Phoneme awareness in language-delayed children: Comparative studies and intervention. *Annals of Dyslexia*, *43*, 153–173.

Weiner, S. (1994). Effects of phonemic awareness training on low- and middle-achieving first graders' phonemic awareness and reading ability. *Journal of Reading Behavior*, *26*, 277–300.

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

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

Wise, B. W., Ring, J., & Olson, R. K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology*, *72*, 271–304.

Wixson, K. K. (1986). Vocabulary instruction and children's comprehension of basal stories. *Reading Research Quarterly*, *21*, 317–329.

Yatvin, J. (2000). Minority View. In National Institute of Child Health and Human Development, *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754, pp. 1–6). Washington, DC: U.S. Government Printing Office.

Appendix: Tables and List of Findings in This Guide

Table 1. Instructional Implications from PA Training Studies Conducted in English and Included in the NRP Report

Student level	*If you want to help them with . . .*	*You could teach them . . .*
Pre-K through first-grade students	Specific phonemic awareness skills	Those skills
First graders	Identifying and reading words	Phonemic awareness skills, using letters, alongside or connected to reading or reading instruction—in other words, teach phonics in context

Table 2. Instructional Implications from Phonics Studies Cited in the NRP Report When Treatment Is Compared to a Basal

If you want to help normally achieving first graders with . . .	*You could . . .*
Better knowing letter names, having higher scores on reading readiness and on a standardized reading test than did children using a basal	Use the New Primary Grades Reading System (Leinhardt & Engel, 1981)
Understanding that reading is a meaning-making process, using phonics more effectively to figure out unknown words, and having greater metacognitive understanding of reading than did children using a basal	Put them in a literature-based/whole language classroom rather than in a traditional/basal-based classroom (Freppon, 1991)
Spelling more unpredictable words and increasing the number of words they use in writing than did children using a basal	Put them in a whole language classroom rather than a traditional classroom (Griffith et al., 1992)

Table 3. Instructional Implications from Phonics Studies (with At-Risk Students) Cited in the NRP Report in Which the Treatment Is Compared to a Basal

For at-risk first graders	
If you want to help at-risk first graders with . . .	*You could . . .*
Vocabulary, comprehension, and phonics (as compared with children whose teacher uses a basal)	Offer explicit phonics instruction as part of whole language instruction (Eldredge, 1991)
Word identification and passage comprehension on the Woodcock-Johnson (as compared to children whose teachers were from the district, used a basal, and were provided with no additional support)	Have master's degree students offer direct instruction and provide university-based supervisors who help the students weekly across the year (Umbach et al., 1989)
Word recognition, nonword reading, and passage reading subtests of the Woodcock-Johnson (as compared to children who read books with their teacher in a small group)	Have graduate students provide 30 minutes of one-on-one Reading Recovery–like tutoring (Santa & Høien, 1999)

Table 4. Instructional Implications from Fluency Studies as Cited in the NRP Report

If you want students to . . .	*You could . . .*
More fluently read a particular text	Have them reread that text (e.g., Neill, 1980, Sindelar et al., 1990)
Better understand a particular text	Have them reread that text (e.g., Sindelar et al., 1990)
	Read the story aloud to them first (e.g., Reitsma, 1988; Rose & Beattie, 1986; Smith, 1979)

Findings from the Comprehension Subgroup Report on Vocabulary

1. Vocabulary should be taught both directly and indirectly.

2. Repetition and multiple exposures to vocabulary items are important.

3. Learning in rich contexts is valuable for vocabulary learning.

4. Vocabulary tasks should be restructured when necessary.

5. Vocabulary learning should entail active engagement in learning tasks.

6. Computer technology can be used to help teach vocabulary.

7. Vocabulary can be acquired through incidental learning.

8. How vocabulary is assessed and evaluated can have differential effects on instruction.

9. Dependence on a single vocabulary instruction method will not result in optimal learning. (Kamil et al., 2000, p. 4-27)

Table 5. Instructional Implications from Comprehension Studies as Cited in the NRP Report

If you want students in grades 3 to 6 to . . .	*You could . . .*
Monitor comprehension	Teach them to monitor (see, for example, Baumann et al., 1992; Block, 1993; Tregaskes & Daines, 1987; Paris et al., 1986)
Know and use a variety of reading strategies, learn more from what they read, and do better on standardized tests	Provide them with teachers who focus instruction on reading strategies across the year (Brown, et al., 1996)
Better remember what they read	Give them graphic organizers (e.g., Alvermann & Boothby, 1986; Darch et al., 1986)
Do better on standardized tests	Have them engage in Reciprocal Teaching in which students are scaffolded into learning how to clarify, question, summarize, and predict (Lysynchuk et al., 1990; Palinscar & Brown, 1984)
Better answer questions about a passage	Teach them ways to find answers (e.g., Raphael & Pearson, 1985; Raphael & McKinney, 1983; Raphael & Wonnacott, 1983; Seretny & Dean, 1986; Garner et al., 1984)
Do better on answering inferential questions	Teach them how to ask inferential questions (Davey & McBride, 1986); teach them cloze procedures, a checklist, or structured overviews (Carr et al., 1991); teach them a reciprocal inference procedure (Reutzel & Hollingsworth, 1991a)
Do better on answering questions about a story	Teach them how to use their knowledge of story structure to ask their own questions (Nolte & Singer, 1985; Short & Ryan, 1984)
Do better on identifying main ideas	Provide direct instruction followed by collaborative practice (Stevens et al., 1987, 1991)
	Provide direct instruction in identifying the main idea (e.g., Rinehart et al., 1986; Carnine et al., 1982)