Ameliorating Early Reading Failure by Integrating the Teaching of Reading and Phonological Skills: The Phonological Linkage Hypothesis

Peter J. Hatcher
Cumbria Education Authority and University of York

Charles Hulme and Andrew W. Ellis
University of York

Hatcher, Peter J.; Hulme, Charles; and Ellis, Andrew W. Ameliorating Early Reading Failure by Integrating the Teaching of Reading and Phonological Skills: The Phonological Linkage Hypothesis. CHILD DEVELOPMENT, 1994, 65, 41-57. We present a longitudinal intervention study of children experiencing difficulties in the early stages of learning to read. Our subjects, 7-year-old poor readers, were divided into 4 matched groups and assigned to 1 of 3 experimental teaching conditions: Reading with Phonology, Reading Alone, Phonology Alone, and a Control. Although the Phonology Alone group showed most improvement on phonological tasks, the Reading with Phonology group made most progress in reading. These results show that interventions to boost phonological skills need to be integrated with the teaching of reading if they are to be maximally effective in improving literacy skills.

There is now a massive body of evidence linking the development of reading skills in children to their underlying phonological skills. This evidence has come from a variety of sources, including studies of children with specific reading difficulties (or dyslexia) and correlational studies of normal children (see Adams, 1990; Goswami & Bryant, 1990; Hulme & Snowling, 1991; and Wagner & Torgesen, 1987, for recent reviews).

The dominant approach to studying the relationship between phonological skills and learning to read has come from studies of phonological awareness. Phonological awareness refers to the ability to reflect explicitly on the sound structure of spoken words. Phonological awareness tasks are among the best predictors of reading skill and, typically, these relationships can be shown to account for significant amounts of variance in reading skill, even after the effects of intelligence have been partialed out (see, e.g., Goswami & Bryant, 1990; and Wagner & Torgesen, 1987, for reviews).

One of the most influential studies of these relationships is that of Bradley and Bryant (1983), who set out to test whether difficulties on one measure of phonological awareness (sound categorization) were caus-

We gratefully acknowledge the help given to us by the teachers and other colleagues (Joan Armstrong, Judith Botham, Deborah Catlett, Pam Chaldecott, Gill Dicken, Janet Hatcher, Penny Kennedy, Mary Kirk [speech therapist], Judith Lord, Cynthia McKerr, Margaret Munn, Joyce Rogers, and Mary Savage) who helped construct the teaching programs, the teachers who participated in the study (Joan Baxter, Richard Burford, Barbara Corbishley, Gill Dicken, Suzanne Edmundson, Claire Ellis, Carol Feeney, Geraldine Gaizely, Mary Gill, Elma Hardy, Anne Harrison, Janet Hatcher, Christine Johnson, Joan Kendall, Anne McGraw, Catherine Rafferty, Ros Rankine, Mary Rice, Joyce Rogers, Ruth Simpson, Gill Slessor, Joyce Smith, and Gordon Weal-leans) and the psychologists (Kathryn Aspinall, Debra Brewer, Peter Crompton, Mike Dawson, Charlie Haigh, Alix Graham, Nigel Harbron, John Heath, Colin McGuinness, Janet Pinney, Don Sinclair, Douglas Thomson, and Michael Watmough), who carried out the greater part of the testing of the children involved. We gratefully acknowledge the willing participation of the children and their parents, the schools and class teachers, and Cumbria Education Authority, without whose support this study would not have been possible. We thank Maggie Snowling for her invaluable help. We also thank three anonymous referees for very helpful comments on an earlier version of this article. Requests for reprints should be sent to: Charles Hulme, Department of Psychology, University of York, York, YO1 5DD, United Kingdom.

[Child Development, 1994, 65, 41-57. © 1994 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920/94/6501-0004/$1.00]
ally related to the development of reading skills. At the beginning of their longitudinal study, the sound categorization ability of over 400 4- and 5-year-old children was assessed before the children started to learn to read. Over 3 years later, their reading and spelling ability and verbal intelligence were assessed. Performance on the sound categorization task was predictive of later reading scores, even when measures of intelligence and memory were taken into account.

In an effort to check that this correlation between early sound categorization skills and reading reflected a causal influence, Bradley and Bryant included a training study. Sixty-five children who initially were poor at sound categorization were split into four groups. One group was trained in sound categorization, and a second group in addition to this was taught letter-sound correspondences (and received exercises relating the sound structure of words to their spelling patterns using plastic letters). There were also two control groups: one was taught to group words according to semantic categories while the other received no training. After training, which was spread over 2 years, the group that had been taught both sound categorization and letter-sound correspondences was some 8–10 months ahead of the semantic categorization control group in reading scores. The group that had only been taught to categorize sounds was about 4 months ahead of the semantic categorization control group in reading; but this difference was not statistically significant. At a later follow-up, 4 years after the first, the same pattern of results was still in evidence, with the group taught sound categorization and letter-sound correspondences still ahead in reading and spelling; though at this point neither trained group was significantly ahead of the taught control group, and interpretation of the data was clouded by ceiling effects (Bradley, 1987). However, given the small numbers of children in this intervention study, these null results must be interpreted with caution: the 4-month difference at first follow-up might well have proved significant given larger numbers of children.

These results are extremely impressive, but they fail to clinch the argument for a causal role of sound categorization ability in learning to read. To prove this would require evidence that the group taught to categorize words only on the basis of their sound was significantly ahead of the group taught to categorize on the basis of meaning. Although there was a tendency for this to happen, the crucial comparison was not statistically reliable.

The difference between the group trained in sound categorization alone and the group who also received training in letter-sound correspondences is notable, however. The exercises that the latter children received involved relating sounds in words to their spelling patterns, in combination with sound categorization training, and led to substantial improvements in reading and spelling skills. A natural implication of this result is that the integration of training in phonological skills with letter-sound training (or more broadly with phonically based reading instruction) may be particularly effective as a way of improving reading skills. According to this view, training in phonological skills in isolation from reading and spelling skills may be much less effective than training that forms explicit links between children’s underlying phonological skills and their experiences in learning to read. We will term this the “phonological linkage hypothesis.”

There are three other important training studies that are consistent with this hypothesis. Lundberg, Frost, and Peterson (1988) trained a large group of Danish kindergarten children in phonological awareness before formal reading instruction began. The trained children were better than a control group at reading in grade 2, though the difference in grade 1 was not quite significant. The size of the effects on literacy skills obtained in this study is small, though such effects may be important, especially if they provide a foundation for further development as children get older. In line with our phonological linkage hypothesis this study does suggest, however, that training on phonological skills alone is not a very powerful way of affecting reading development.

In the study of Cunningham (1990), two groups of kindergarten and first-grade children were given different forms of phonemic awareness training. A “skill and drill” group received training in phoneme segmentation and blending, while a “metalevel” group received training that in addition explicitly emphasized the link between the phonemic awareness training and reading. This latter training involved exercises in which the applications of segmentation and blending in reading were demonstrated and practiced. A control group listened to stories and discussed them with a teacher.

Cunningham found that the trained
groups made more progress with reading than the control group and that, among the older children, the “metalevel” training had more effect than “skill and drill” training. This result fits nicely with our phonological linkage hypothesis. However, as in the case of Bradley and Bryant’s sound categorization and letter-sound group, the “metalevel” group really received both teaching in reading and phonemic awareness training. Without a group simply given reading instruction alone, it is hard to be sure that it really is the integration of phonological training and reading instruction that is crucial to this group’s success.

The third training study is that of Ball and Blachman (1988, 1991). Their study involved two groups of kindergarten children. A “phoneme awareness” training group received training in word segmentation, letter names and sounds, sound categorization (as in Bradley and Bryant’s study), and DISTAR spell-by-sounds training. A “language activities” group received training in letter names and letter sounds and general language activities. There was also an unseen control group. The results showed that reading and spelling scores improved most in the “phoneme awareness” group. In line with our phonological linkage hypothesis, this study shows that phonological training combined with the teaching of letter names, letter sounds, and spelling skills is effective.

As well as these training studies, there is evidence from a number of experimental studies that is consistent with our phonological linkage hypothesis. For example, Tunmer, Herriman, and Nesdale (1988) looked at the relation between phonemic awareness and nonword reading (a relatively pure measure of decoding skill). They found that, although all the children who performed well on the nonword decoding task had good phonemic awareness skills, there were some children who had good phonemic awareness but read nonwords poorly. They conclude that “phonological awareness is necessary but not sufficient for acquiring phonological recoding skill.” A similar conclusion was drawn by Byrne and Fielding-Barnsley (1989). They looked at young, preliterate children’s ability to understand the alphabetic principle; that is, to demonstrate an understanding that particular phonemes in words are represented systematically by particular letters. To examine this they first taught 3–5-year-old children to read the words “mat” and “sat.” The children were then simply asked to decide whether the printed word “MOW” was pronounced as “mow” or “sow.” Above-chance performance on this task requires an understanding of the alphabetic principle while placing minimal extra cognitive demands on very young children. Byrne and Fielding-Barnsley found that reliable performance on this task was achieved only by children who could perform phonemic segmentation, understood that the initial sound segments of different words shared the same identity (that /s/ in SAT is the same as /s/ in SOW, e.g.), and had also learned the letters corresponding to the sounds ‘m’ and ‘s’. Thus both phoneme awareness and knowledge of letter identity were needed for children to grasp the alphabetic principle which is one component of a phonic decoding reading strategy: neither phoneme awareness nor letter identity knowledge alone was sufficient.

The phonological linkage hypothesis has important implications for our understanding of the way children learn to read and also has important educational implications. The main aim of the present study is to test this hypothesis in the context of an educationally realistic longitudinal intervention study with children who are showing significant difficulties in the early stages of learning to read.

In the present study we chose to look at the effectiveness of structured interventions in alleviating the reading problems of a large sample of poor readers. We chose to study 7-year-olds because by this age it is possible to identify with some certainty those children who are experiencing difficulties in learning to read. Following from the phonological linkage hypothesis, we wished to test whether an intervention that involved a combination of phonological training and reading instruction would be more effective than an intervention involving either reading instruction alone or phonological training alone.

This design allows us to assess whether simply training phonological skills is enough to improve the reading skills of poor readers: in this case the reading of both groups given phonological training should improve. In contrast, our phonological linkage hypothesis makes the specific prediction that the children given an integrated combination of reading and phonological training should make more progress than any of the other groups. Support for our hypothesis would have obvious practical implications
for the design of teaching programs for children experiencing difficulties in the early stages of learning to read.

**Method**

In the present study we examine the effectiveness of three different structured methods of teaching with children who are having difficulties in the early stages of learning to read. We also relate progress made by these children to measures of their phonological skills taken before teaching began. The children were assessed before teaching began (t1, in September 1989) and after teaching was completed (t2, April–May 1990) some 7 months later. The 20 weeks of teaching were spread over a 25-week period between mid-October and early April. To assess the durability of any effects on reading and spelling, all children were reassessed 9 months after the interventions had ceased (t3, January 1991). All assessments were carried out "blind" as to the group membership of the child.

**Subjects**

The starting point for the present study was a county-wide reading survey of 6- and 7-year-olds in their third year of infant schooling in Cumbria Education Authority, United Kingdom. This screening used the Carver (1970) test, a group-administered, single word reading test, where children have to underline one of a group of words to match the word that has been spoken by the examiner. Following the survey, 188 children were identified as having reading quotients of less than 86. These children were then screened for severe general learning difficulties using the Raven's Coloured Progressive Matrices (Raven, 1965). Twenty children with a percentile rank on the matrices of less than 25 and a Carver reading quotient of less than 71 were excluded. Other children were excluded for a wide variety of reasons including failure to obtain parental consent to participate (six), changes in school (nine), and children being given Statements of Special Educational Need which give them a legal entitlement to special educational provision (seven). From the remaining children, 128 were selected and divided into four groups of 32, matched on IQ, reading ability, and age. Subsequently, however, three children were lost from the study, as they moved out of the area, reducing the total sample to 125. The children included in the study showed a wide range of IQs (68–122) and can be considered representative of 7-year-old children experiencing reading problems.

For the purposes of matching subject groups, two verbal (similarities and vocabulary) and two performance (object assembly and block design) subtests of the WISC-R (Wechsler, 1974) were used to estimate IQ. Scaled scores were converted to short-form IQ scores (with a mean of 100 and a standard deviation of 15) according to the procedure proposed by Tellegen and Briggs (1967; see also Sattler, 1982). Matching on reading ability was achieved using the British Ability Scales (BAS) Word Reading Test (Elliott, Murray, & Pearson, 1983). On the basis of these scores the four groups were closely matched for sex, age, IQ, and reading age. The distribution of children between groups was balanced within and across schools. There were 18 boys and 14 girls in each group. The four groups were assigned to one of three experimental teaching conditions and a control condition. These were Reading with Phonology (R+P), Reading Alone (R), Phonology Alone (P), and a Control (C).

The details of the four groups are shown in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEANS (and Standard Deviations) for Age, BAS Word Reading Ages, and Prorated WISC-R Full-Scale IQ for the Four Groups (N = 124)</strong></td>
</tr>
<tr>
<td>Group</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Reading with Phonology (N = 32)</td>
</tr>
<tr>
<td>Reading Alone (N = 31)</td>
</tr>
<tr>
<td>Phonology Alone (N = 30)</td>
</tr>
<tr>
<td>Control (N = 31)</td>
</tr>
<tr>
<td>F(3, 120) ratio</td>
</tr>
</tbody>
</table>
PROCEDURE

Cognitive Measures

At the beginning of the project, all children received a large battery of 19 tests measuring general intellectual, reading, spelling, arithmetic, memory, and phonological skills. These tests were administered in a fixed order in three sessions. We describe here only the subset of 12 measures analyzed in the present paper.

Reading.—Progress in reading was assessed by tests of word and nonword pronunciation and text reading accuracy and comprehension.

Early Word Recognition Test: An Early Word Recognition Test was constructed to differentiate between the progress of children who were at a very early stage of acquiring a "sight vocabulary." Forty-two words found at the first book stage of seven different reading schemes used in the schools participating in the study were used. Scoring was based upon the total number of words read aloud correctly.

British Ability Scales (BAS) Word Reading Test A: This test (Elliott et al., 1983) was used as a standardized measure of context-free word recognition. The test consists of a graded list of 90 words and was administered before matching groups and again after the intervention was completed.

Neale Analysis of Reading Ability (revised, form 1): This test (Neale, 1989) was used as a measure of reading accuracy in context and of reading comprehension. The test comprises a graded sequence of passages, each of which has an accompanying list of comprehension questions.

Nonword Reading Test: This test was included as a relatively pure measure of phonetic decoding skill. A list of 70 phonetically legal nonwords was presented to each child to pronounce. The test was graded in difficulty from simple one-syllable CVC (consonant, vowel, consonant) nonwords to complex multisyllabic forms. It was explained to the child that the test consisted of "pretend" words that they would not know and that they should attempt to say as many of the "pretend" words as they could. Testing was discontinued when the child had failed on five consecutive items. Scoring was based on the number of nonwords read correctly with a phonically regular pronunciation.

Spelling.—The Schonell (Schonell & Schonell, 1956) Graded Word Spelling Test: List B was used to measure spelling ability. This test consists of 100 words (mostly phonically regular) which range from CVC to multisyllabic word forms.

Arithmetic.—The BAS Basic Number Skills Test: Test A (Elliott et al., 1983) was used to measure arithmetic skills. The test consists of 34 orally presented items relating to a booklet of pictures and 34 paper and pencil problems of graded difficulty.

Phonological skills.—Sound deletion: A modified version of Bruce's (1964) Word Analysis Test was used to measure the ability to delete sounds from spoken words. There were 24 items, six in each section of the test. Before presenting the test, the examiner introduced the concept of sounds in words by using examples such as detecting the sound /k/ in 'cat'. Each of the first three test sections was preceded by two practice items which were presented according to a similar format. When introducing the first section (deletion of beginning sounds), the examiner said, "Say the word 'seat'. . . What word is left if the /s/ sound is taken away from the beginning of 'seat'?" Where children could not do that they were told, "If we take the /s/ sound away from 'seat' we are left with the word 'eat'.” The examiner presented the second example for section 1 by asking, "What word is left if the /p/ sound is taken away from the beginning of 'pin'?" Where necessary help was given. The same format was used when presenting examples for sections 2 (deletion of final sounds) and 3 (deletion of middle sounds). The examiner presented the questions for sections 1–3 by using the stem, "What word is left if . . . ?" The first two items in each section continued with reference to the sound and position of elision, for example, " . . . the /j/ sound is taken away from the beginning of 'jam'?") The next two items continued with reference to just the sound to be deleted, for example, " . . . the /s/ sound is taken away from 'spin'.” The last two items continued without reference to either the concept “sound” or to the position of elision, for example, " . . . the /h/ is taken away from 'hill'.”

Section 4 was preceded by the examiner’s remarking, “The sounds that I want you to listen to now will be in different positions in the words.” Section 4 items were presented using the stem, "What word is left if . . . ?" and without reference to either the concept “sound” or to the position of elision which varied from item to item.

Section 1 was discontinued at the point where children made five consecutive er-
rors. The entire test was discontinued at the point where children made five consecutive errors beyond the first item of section 2. Scoring was based upon the overall number of correct responses.

**Sound blending:** A Sound Blending Test was constructed and used to measure the ability to blend a sequence of sounds into nonwords. The test stimuli consisted of 30 sets of two to seven sounds that could be blended to produce nonwords. In the order of presentation there were five sets of two-sound (VC or CV), six sets of three-sound (CVC), eight sets of four-sound (CVCC or CCVC forming final and initial blends), five sets of five-sound (VCVCC or VCCVC), and five sets of six-sound combinations (CVCCVC or VCCCVCC) and one seven-sound combination (CVCCCVC).

Each set of sounds was read by the examiner at the rate of two sounds per second. The sounds were either consonants or short vowels. When presenting consonant sounds, the examiner tried to ensure that the vowel following each consonant was reduced to a minimum. The test was discontinued after children had made five consecutive errors. Scoring was based upon the number of correct responses.

**Nonword segmentation:** The test stimuli consisted of 30 sets of nonwords, each of which consisted of from two to seven phonemes. In the order of presentation there were five items with two sounds (VC or CV), seven with three sounds (CVC), seven with four sounds (CVCC or CCVC), five with five sounds (VCVCC or VCCVC), four with six sounds (VCCVCC, CCVCCV, or CVCCVC), and two with seven sounds (CVCCVCC and VCCCVCC).

The examiner introduced the task by saying, "A nonword is a word like 'ot' or 'ip'. It sounds a bit like a real word but it doesn't make sense. It is a pretend word. 'Uk' and 'af' are also nonwords." Two coins were put on the table, and the examiner introduced the first of three examples by saying, "I am going to say a nonword. I will then use these coins to break it up into separate sounds." After saying 'ot' the examiner pronounced the sounds /o/-/t/ and simultaneously with each sound pushed a coin forward. The examiner then said, "Did you see how I left a gap between each sound? Now I want you to say the nonword 'ip' and to break it up into separate sounds just like I did. Use the coins and leave a gap between each sound." If necessary, the task was demonstrated and children asked to copy what they had been shown. Two more examples were given following the same format. In the examples used to introduce the test, the examiner pronounced the separate sounds at the rate of two per second. The vowel sound after each consonant was minimized; all vowels were short. The examiner introduced test items by saying, "Now I am going to say some more nonwords. Use the coins to break them up into separate sounds. Only say the sounds that you hear. Leave a gap between each sound." For each successive item the number of coins corresponding to its constituent sounds was put on the table and the nonword pronounced.

Children's responses were accepted as correct when they produced the appropriate sounds, each of them separated in time from other sounds and accompanied by the simultaneous pushing forward of a coin. The test was discontinued after children had made five consecutive errors. Scoring was based on the number of correct responses.

**Sound categorization:** A modified version of Bradley's (1984) sound categorization test was used to measure the ability to recognize rhyme and alliteration in spoken words. The experimental stimuli were 30 sets of four words. Within each set three words contained a common sound that the fourth lacked. The first 20 sets were rhyme oddity tasks, with the distinctive sound being the last consonant in the first 10 sets and the medial vowel in the second 10. The final 10 sets of words constituted an alliteration oddity task. Each set of tasks was preceded by two examples or practice items. Scoring was based upon the total number of correct responses.

**Teaching Procedures**

Our aim was to contrast the effectiveness, in boosting our subjects' reading skills, of an integrated program of reading with phonological training (R+P), with training in phonology alone (P), or with reading instruction alone (R). Each program involved the children being taught individually for 40 30-min sessions spread over 20 weeks. The children in the three experimental groups were taught by a total of 23 teachers; each teacher worked individually with a total of between two and nine children. Since most of the teachers were involved with sets of three children (one from each experimental condition), the times of the day that the children were taught was varied. The minimum period of time between any one child's
teaching sessions was 42 hours. All teachers received 3 days' training in how to use the teaching materials prepared for each of the interventions. The teachers involved in the research project comprised eight learning-support teachers who visited the schools, seven school-based special-needs teachers, seven class teachers, and one head teacher. The teachers were granted relief from their normal duties in order to carry out the research interventions. Adherence to the teaching protocols was monitored via regular meetings with the principal investigator and the completion of written records for each teaching session.

For reasons of space we give here only brief details of the teaching procedures used. (Further details of the procedures, together with samples of the materials used, are available from the first author.)

Phonological Training Alone (P).—This training package (which was purely phonological, involving no reading) consisted of nine sections of phonological tasks based broadly on the “levels of difficulty” and activities referred to by Lewkowicz (1980), Lundberg et al. (1988), Rosner (1975), Stanovich, Cunningham, and Cramer (1984), and Yopp (1988). The sections covered the identification and supply of rhyming words, the identification of words as units within sentences, the identification and manipulation of syllables, the identification and discrimination of sounds within words, sound synthesis (into words), word segmentation (into sounds), the omission of sounds from words, sound substitution within words, and the transposition of sounds within words. Some of the activities (sound deletion, word segmentation, and sound synthesis) were equivalent to those used in the phonological portions of the assessment battery, in that they involved the same tasks but different stimuli. Each section consisted of a number of activities (68 in the package) which varied in terms of teaching mode, cognitive task, and level of stimulus difficulty. The package was ordered, in terms of progressing from easier to more difficult activities.

The activities were presented as sequenced in the package and at a rate commensurate with children’s success with the materials. Teachers were provided with record sheets with which to monitor children’s progress. At the end of each 30-min session teachers entered the child's score on the record sheet appropriate for each of the activities covered in the session. A criterion of at least 80% success was used to determine when a child moved from one activity to the next or from one section to the next.

Reading with Phonology (R+P).—This package was modeled on the work of Clay (1985) but included the addition of phonological activities from the phonology (P) condition described above. Several changes were made in the way that Clay’s (1985) teaching strategies and diagnostic tests were presented. The individual tests and teaching strategies were set out in sections covering reading and writing in context, reading and writing words, and looking at and writing letters. Each of the sections had a common format with the aims, materials needed, and instructions specified. Each of the teaching strategies also had specified criteria for their use and discontinuation. Identified areas of weakness were linked to specific teaching strategies by a chart.

The first four teaching sessions were taken up with the assessment of each child’s responses to easy and to hard text, words, and letters. Sessions 5–40 followed the teaching format outlined by Clay (1985). This included re-reading a book that could be read with greater than 94% accuracy, reading the book introduced at the end of the previous session with the teacher taking a running record, letter identification (where necessary), writing a story (linked to phonological activities), cutting up the story (where necessary), and introducing a new book and attempting to read it. The criterion for moving up to the next level of difficulty was greater than 94% reading accuracy, over three sessions. A common set of 73 books, split into 20 levels (graded in difficulty according to the list provided by the New Zealand Department of Education, 1987) was used by all teachers. There were three or four books at each difficulty level.

Some comparisons between the amount of time spent on phonological training in the P and R+P groups can be made. The P group obviously spent more time on the phonological activities and completed on average 66.97 of the 68 phonology activities. In the R+P group the children completed a minimum of 31.21 of the phonological activities. In addition, however, the children in the R+P group carried out activities to link reading and phonology that were not included in the P group’s activities. These linking activities included practicing letter-sound associations, relating spellings to sounds using plastic letters (as advocated by
Bradley & Bryant, 1983), and writing words while paying attention to letter-sound relationships. The phonological and phonological linkage activities were carried out in the middle part of each session and lasted approximately 10 min. The phonological linkage exercises were linked to points arising from the children's writing (writing a story) and their reading during the first part of each session.

**Reading Alone (R).**—The Reading Alone package was identical to that used with group R+P except for the omission of any explicit reference to phonology and all of the teaching strategies explicitly concerned with phonological linkage activities. In the early stages of the intervention period, teachers were frequently reminded of the importance of not referring to phonology or letter-sound relationships when working in this condition. Where children already exhibited such skills, teachers accepted them but made neither positive nor negative comments about them.

Some comparisons between the amount of time spent on reading instruction, and the form it took, in the R and R+P groups can be made. In the R group more time was spent on reading books of appropriate difficulty and building up reading and writing vocabularies through the use of structured multisensory teaching techniques (as advocated by Bryant & Bradley, 1985) in place of the phonological and phonological linkage activities undertaken by the R+P group. Though these children were not taught letter-sound associations, they were taught letter names. In addition, their reading instruction devoted more time to teaching the usefulness of context and meaning in reading and the use of self-checking and correction for attempts at reading unknown words. In short, this program involved individualized, highly structured teaching, embodying many current recommendations for good practice (Pumfrey & Elliott, 1990), but it lacked the explicit phonological linkage instruction given to the R+P group. It should be noted, however, that questionnaires completed by the children's classroom teachers indicated that virtually all the children in the study were receiving phonic reading instruction at school. The extent of this instruction showed no systematic differences between the different groups in the study. In no cases did such phonic teaching involve exercises in phonological awareness of the sort used for the P and R+P groups.

**Control (C).**—These children received their regular classroom teaching without any special form of additional provision from our study. Some of these children, however, like those in the other groups, were receiving additional remedial teaching that was independent of that provided by our study. Fifteen children in the control group received such help compared to five in the R+P group, seven in the R group, and six in the P group. Clearly, these differences in the amount of independent help given tend to operate against our hypothesis, since the largest number of children receiving help was in the control group and the smallest number was in the R+P group.

**Results**

Our aim in this study was to assess the differential effectiveness of three teaching methods in helping children who are experiencing difficulties in the early stages of learning to read. Our primary data, therefore, come from the changes in reading (and also spelling) between the pre- and posttests. The inclusion of arithmetic provides an important control capable of showing that any gains are specific to the domain of literacy rather than reflecting general, nonspecific improvements. We present the data on attainments first before going on to deal with the results from the phonological tasks.

**ATTAINMENT IN READING, SPELLING, AND ARITHMETIC**

The means (and standard deviations) for the reading, spelling, and arithmetic measures on the pre- and posttests for the four groups are shown in Table 2 below. The results of the standardized tests are presented in terms of attainment ages since these are easy to interpret, though in all cases analyses were conducted on the raw scores.

The children in the four groups were matched for reading ability using the BAS Word Reading Test only. As can be seen from Table 2, although the groups are closely matched on BAS reading age, they are not as closely matched as we would have hoped on the other measures of literacy skill.

The main issue of interest is the extent to which the four groups have made differential progress on these attainment measures. Analyses to address this question need to take account of the fact that the groups are not perfectly matched for their levels of literacy attainment at t1. The t1 literacy scores (Early Word Recognition, BAS
TABLE 2
MEANS (and Standard Deviations) FOR THE PRE- AND POSTINTERVENTION ATTAINMENT MEASURES OF READING, SPELLING, AND ARITHMETIC IN THE FOUR GROUPS

<table>
<thead>
<tr>
<th>GROUP</th>
<th>Reading with Phonology (N = 32)</th>
<th>Reading Alone (N = 31)</th>
<th>Phonology Alone (N = 30)</th>
<th>Control (N = 31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Word Identification:a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>20.22 (10.08)</td>
<td>20.10 (9.49)</td>
<td>21.03 (11.63)</td>
<td>20.90 (9.59)</td>
</tr>
<tr>
<td>t2</td>
<td>32.72 (10.55)</td>
<td>32.32 (7.68)</td>
<td>29.73 (10.54)</td>
<td>29.32 (9.05)</td>
</tr>
<tr>
<td>BAS Word Reading:b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>5.85 (.53)</td>
<td>5.90 (.47)</td>
<td>5.90 (.57)</td>
<td>5.96 (.53)</td>
</tr>
<tr>
<td>t2</td>
<td>6.73 (.85)</td>
<td>6.56 (.43)</td>
<td>6.55 (.69)</td>
<td>6.60 (.67)</td>
</tr>
<tr>
<td>Neale Accuracy:b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>5.10 (.21)</td>
<td>5.04 (.19)</td>
<td>5.18 (.43)</td>
<td>5.11 (.30)</td>
</tr>
<tr>
<td>t2</td>
<td>6.13 (1.00)</td>
<td>5.75 (.54)</td>
<td>5.81 (.90)</td>
<td>5.66 (.80)</td>
</tr>
<tr>
<td>t3</td>
<td>6.77 (1.88)</td>
<td>6.22 (.82)</td>
<td>6.31 (1.03)</td>
<td>6.25 (1.15)</td>
</tr>
<tr>
<td>Neale Comprehension:b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>5.29 (.30)</td>
<td>5.32 (.34)</td>
<td>5.43 (.50)</td>
<td>5.41 (.49)</td>
</tr>
<tr>
<td>t2</td>
<td>6.39 (.92)</td>
<td>6.00 (.97)</td>
<td>5.94 (.80)</td>
<td>5.88 (.73)</td>
</tr>
<tr>
<td>t3</td>
<td>6.99 (1.28)</td>
<td>6.47 (.94)</td>
<td>6.46 (1.11)</td>
<td>6.35 (.97)</td>
</tr>
<tr>
<td>Nonword Reading:c</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>4.34 (4.53)</td>
<td>3.55 (3.71)</td>
<td>6.00 (7.28)</td>
<td>3.65 (5.51)</td>
</tr>
<tr>
<td>t2</td>
<td>15.39 (14.16)</td>
<td>10.77 (8.14)</td>
<td>15.53 (10.28)</td>
<td>11.87 (10.97)</td>
</tr>
<tr>
<td>Schonell Spelling:b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>5.78 (.59)</td>
<td>5.83 (.50)</td>
<td>5.93 (.56)</td>
<td>5.77 (.55)</td>
</tr>
<tr>
<td>t2</td>
<td>6.77 (.93)</td>
<td>6.54 (.55)</td>
<td>6.66 (.63)</td>
<td>6.49 (.74)</td>
</tr>
<tr>
<td>t3</td>
<td>7.19 (1.02)</td>
<td>6.90 (.62)</td>
<td>6.99 (.82)</td>
<td>6.92 (.78)</td>
</tr>
<tr>
<td>BAS Arithmetic:b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>6.64 (.61)</td>
<td>6.83 (.64)</td>
<td>6.76 (.57)</td>
<td>6.69 (.66)</td>
</tr>
<tr>
<td>t2</td>
<td>7.39 (.59)</td>
<td>7.50 (.51)</td>
<td>7.37 (.63)</td>
<td>7.44 (.57)</td>
</tr>
</tbody>
</table>

*a* Maximum score = 42.
*b* Attainment ages expressed in years.
*c* Maximum score = 70.

Word Reading, Neale Accuracy, Neale Comprehension, Nonword Reading, and Spelling) for all subjects were subjected to a principal components analysis. This yielded a single factor (literacy skill) accounting for 70.1% of the variance. In subsequent analyses, the t1 literacy skill factor scores were used as a covariate when examining differences between groups at t2 on the different measures of literacy attainment. In each analysis, planned contrasts were used to test whether each of the trained groups differed from the control group. The use of planned contrasts does not require a significant overall effect for groups, and we therefore present only the results of the contrasts. Our prediction was that the R+P group would be consistently ahead of the control group in literacy skills at t2.

We chose to present separate analyses of our different measures of literacy skill (rather than a single multivariate analysis of covariance) because these measures were chosen to give an assessment of partially independent components of literacy skill. The Early Word Recognition and BAS tests provide a measure of context-free word recognition skill. The Neale Analysis of Reading Ability measures word recognition ability in context and the ability to comprehend what has been read. The Nonword Reading Test provides a relatively pure measure of children's ability to apply phonic decoding strategies to unfamiliar words out of context. The Schonell spelling test provides a measure of spelling ability. We wished to assess the extent to which our interventions had affected each of these measures; these measures are, however, correlated, and there is a degree of redundancy in these separate analyses (Huberty & Morris, 1989).

The use of analysis of covariance requires that the covariate is correlated with the dependent variable and that there is homogeneity of regression between the covariate and the dependent variable in the dif-
ferent groups considered (Tabachnick & Fidell, 1989). Both of these assumptions were met by our data for the literacy and arithmetic measures. Prior to conducting analyses, the data were also screened for outliers (Tabachnick & Fidell, 1989). One child in Group R was clearly an outlier as his scores on two of the t2 literacy measures (BAS and Neale Accuracy) were more than 3.5 SD above the group mean. This child was therefore excluded from all analyses of the data, reducing the total sample to 124.

Reading

As can be seen from Table 2, the improvements in reading following the intervention, at t2, tend to be consistently larger in the Reading with Phonology (R+P) group.

For the Early Word Recognition test, Group R+P differed significantly from the Control group, $F(1, 119) = 7.90, p < .01$, as did Group R, $F(1, 119) = 7.17, p < .01$, though Group P, $F(1, 119) = 0.23, N.S.$, did not differ significantly from the Control.

For the BAS word recognition test, Group R+P differed significantly from the Control group, $F(1, 119) = 5.90, p < .02$, but neither Group R, $F(1, 119) = 0.26, N.S.$, nor Group P, $F(1, 119) = 1.21, N.S.$, differed significantly from the Control.

For the Neale Analysis of Reading Ability Accuracy scores, Group R+P differed significantly from the Control group, $F(1, 119) = 15.97, p < .001$, but neither Group R, $F(1, 119) = 2.96, N.S.$, nor Group P, $F(1, 119) = 0.09, N.S.$, differed significantly from the Control.

For the Neale Analysis of Reading Ability Comprehension scores, Group R+P differed significantly from the Control group, $F(1, 119) = 11.95, p < .001$, but neither Group R, $F(1, 119) = 1.33, N.S.$, nor Group P, $F(1, 119) = 0.06, N.S.$, differed significantly from the Control.

For the Nonword Reading scores, Group R+P differed significantly from the Control group, $F(1, 119) = 4.00, p < .05$, but neither Group R, $F(1, 119) = 0.02, N.S.$, nor Group P, $F(1, 119) = 1.45, N.S.$, differed significantly from the Control.

Thus, for every one of our reading measures, the R+P group has made significantly more progress than the control group, and in every case apart from one, the other treated groups have failed to make significantly more progress than the control group.

Spelling

As in the case of reading, it appears that Group R+P has made more progress in spelling than the other treated groups. The spelling scores were analyzed in the same way as the reading scores; Group R+P differed significantly from the Control group, $F(1, 119) = 5.88, p < .02$, but neither Group R, $F(1, 119) = 0.68, N.S.$, nor Group P, $F(1, 119) = 0.49, N.S.$, differed significantly from the Control.

Arithmetic

In contrast to the results for reading, the changes in arithmetic skills seem similar in all four groups. An analysis of covariance was conducted on the t2 arithmetic scores with the t1 scores as the covariate, followed by planned contrasts, as for the reading and spelling measures. This showed that in no case did any of the treatment groups differ from the control (Group R+P, $F(1, 119) = 0.04, N.S.$, Group R, $F(1, 119) = 0.12, N.S.$, Group P, $F(1, 119) = 0.83, N.S.$). This pattern confirms that the previous differential effects observed for reading and spelling are not a consequence of any general, artifactual improvement in Group R+P.

LONG-TERM EFFECTS OF INTERVENTION: READING AND SPELLING SCORES 9 MONTHS LATER

Reading

Reading was assessed using the Neale Analysis, and spelling with the Schonell test, 9 months after the intervention had ceased. Table 2 shows that the larger improvements in reading on the Neale in the R+P group appear to be maintained at t3, 9 months after teaching finished. An analysis of the t3 scores, using the t1 literacy skill factor scores as a covariate, followed by planned contrasts, was conducted.

For the Neale Analysis of Reading Ability accuracy scores, Group R+P differed significantly from the control group, $F(1, 119) = 8.48, p < .01$, but neither Group R, $F(1, 119) = 0.29, N.S.$, nor Group P, $F(1, 119) = 0.15, N.S.$, differed significantly from the Control. An identical pattern emerged for the Neale Analysis of Reading Ability comprehension scores, where Group R+P differed significantly from the Control group, $F(1, 119) = 12.38, p < .001$, but neither Group R, $F(1, 119) = 1.06, N.S.$, nor Group P, $F(1, 119) = 0.01, N.S.$, differed significantly from the Control.

Thus the improvements in reading skill shown by the group given the integrated
phonological and reading package were maintained 9 months after our intervention had ceased.

Spelling

It appears from Table 2 that the differential effect on spelling at t2 has become weaker by t3. An analysis of the t3 spelling scores using the t1 literacy skill factor scores as a covariate followed by planned contrasts showed that in no case did any of the treatment groups differ from the control, Group R + P, \( F(1, 119) = 2.46 \), N.S., Group R, \( F(1, 119) = 0.06 \), N.S., Group P, \( F(1, 119) = 0.01 \), N.S.

Changes in Phonological Skills

Our phonological linkage theory holds that in order to be effective in boosting reading skills the training of phonological and reading skills needs to be integrated. Our results support this position insofar as the group given integrated reading and phonological training (R + P) improved more in reading skills than did the other groups who were given equal amounts of teaching concentrated solely on reading (R) or on phonological training (P). Moreover, the improvements seen were selective to reading, since a similar pattern of gains was not seen in arithmetic. The gains in reading were also shown to be durable.

One possible objection however, would be that the integration of reading and phonological training was more effective because (contrary to a naive view) this package was actually more effective in improving phonological skills than was the phonological training alone. According to this argument there might be something uniquely effective for improving phonological skills in combining phonological exercises with explicit reference to the printed word as happened in the R + P group. Such a view is quite plausible. Given the highly abstract nature of the phoneme as a unit of speech (see, e.g., Liberman, Shankweiler, Fischer, & Carter, 1974) it could be that children would benefit greatly in learning to perform tasks involving phoneme manipulation from the availability of visual letter symbols which stand in a direct correspondence to them. On this view, the greater improvements in literacy skills seen in the R + P group may, in reality, be a simple product of the better training of phonological skills in this group. There is an obvious way in which to test this idea: since we have independent measures of phonological ability taken before and after the teaching interventions we can compare the groups on these measures.

The means (and standard deviations) for the four phonological measures on the pre- and posttests for each of the four groups are shown in Table 3. As can be seen, there are substantial improvements in phonological skills in all of the groups, but the size of these improvements is consistently larger in the Phonology (P) group. To examine the reliability of any differential improvement an

### Table 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Reading with Phonology ( (N = 32) )</th>
<th>Reading Alone ( (N = 31) )</th>
<th>Phonology Alone ( (N = 30) )</th>
<th>Control ( (N = 31) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound deletion:*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>2.88 (3.08)</td>
<td>2.94 (3.80)</td>
<td>3.80 (4.27)</td>
<td>2.39 (3.05)</td>
</tr>
<tr>
<td>t2</td>
<td>9.91 (7.41)</td>
<td>6.19 (4.56)</td>
<td>13.70 (6.17)</td>
<td>7.39 (6.20)</td>
</tr>
<tr>
<td>Nonword segmentation:*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>9.31 (8.04)</td>
<td>10.29 (6.89)</td>
<td>11.40 (7.19)</td>
<td>8.32 (7.31)</td>
</tr>
<tr>
<td>t2</td>
<td>15.81 (8.60)</td>
<td>16.10 (6.88)</td>
<td>18.63 (6.09)</td>
<td>14.94 (7.66)</td>
</tr>
<tr>
<td>Sound blending:*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>6.91 (4.94)</td>
<td>5.97 (4.67)</td>
<td>7.57 (5.81)</td>
<td>7.13 (6.12)</td>
</tr>
<tr>
<td>t2</td>
<td>12.09 (5.52)</td>
<td>10.71 (5.29)</td>
<td>14.43 (5.16)</td>
<td>11.03 (7.11)</td>
</tr>
<tr>
<td>Sound categorization:*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>14.31 (6.06)</td>
<td>13.81 (5.64)</td>
<td>14.60 (5.61)</td>
<td>13.29 (4.71)</td>
</tr>
<tr>
<td>t2</td>
<td>17.56 (5.77)</td>
<td>16.90 (5.85)</td>
<td>18.27 (5.85)</td>
<td>15.65 (5.42)</td>
</tr>
</tbody>
</table>

* Maximum score = 24.

* Maximum score = 30.
analysis of variance was carried out on the post- \((t2)\) minus pre-intervention \((t1)\) difference scores followed by planned contrasts comparing each of the treatment groups with the control. (In this case we were unable to use analyses of covariance, analogous to the ones used for reading and spelling, because tests revealed heterogeneity of regression across groups between two of the phonological measures [sound deletion and nonword segmentation] and phonological ability factor scores derived from a principal components analysis.)

For the Sound Deletion test, Group P, \(F(1, 120) = 11.14, p < .001\), differed significantly from the Control group, but neither Group R, \(F(1, 120) = 1.43, \text{N.S.}\), nor Group R + P, \(F(1, 120) = 1.98, \text{N.S.}\), differed significantly from the Control. Similarly, for the Sound Blending test, Group P differed significantly from the Control group, \(F(1, 120) = 6.61, p < .01\), but neither Group R, \(F(1, 120) = 0.54, \text{N.S.}\), nor Group R + P, \(F(1, 120) = 1.28, \text{N.S.}\), differed significantly from the Control. However, for Sound Categorization in no case did a treatment group differ from the control, Group P, \(F(1, 120) = 0.92, \text{N.S.}\), Group R, \(F(1, 120) = 0.30, \text{N.S.}\), Group R + P, \(F(1, 120) = 0.44, \text{N.S.}\), and the same was true for Nonword Segmentation, Group P, \(F(1, 120) = 0.20, \text{N.S.}\), Group R, \(F(1, 120) = 0.01, \text{N.S.}\), and the Control.

There is a clear trend for Group P to make more progress on the phonological measures than the other groups, although in only two cases is this group’s progress significantly better than that of the Control group. Given this pattern, it seemed possible that by looking at the joint effects of these measures in a single analysis we would find an overall superiority of Group P to the Control group. In some cases although individual analyses of variance are not significant, a multivariate analysis that takes account of the joint effect of a number of related variables can produce a significant effect (Tabachnick & Fidell, 1989). To explore this possibility, the difference scores \((t2 - t1)\) for the four phonological measures were entered into a multivariate analysis of variance followed by planned contrasts. These revealed that, on a composite measure of phonological skill, Group P had made significantly more progress than the Control group, \(F(4, 117) = 4.20, p < .005\), but neither Group R, \(F(4, 117) = 0.74, \text{N.S.}\), nor Group R + P, \(F(4, 117) = 0.86, \text{N.S.}\), differed significantly from the Control. Overall, therefore, Group P has made significantly more progress in phonological skills than the Control group while neither of the other two treated groups has.

These results show clearly that the differences in literacy scores between the P and R + P groups cannot simply be attributed to differences in the extent of improvements in phonological skills. Only Group P made significant progress in phonological skills, but this was not matched by comparable improvements in their literacy skills. In contrast, the R + P group did not make significant improvements in phonological skills, but they were the only group to make significant progress in learning to read and to spell. These results also suggest that it is relatively difficult to improve phonological skills in poor readers. Although the R + P group spent an appreciable amount of time being trained on phonological tasks, this apparently was not sufficient to bring about a significant improvement in their phonological skills.

### Discussion

This longitudinal intervention study of children with difficulties mastering early reading skills has produced a number of findings of both practical and theoretical significance. Before discussing our results it is important to emphasize that the intervention we conducted did not involve total control of the children’s experience in learning to read. Rather our intervention involved just a small supplement of individualized tuition that was additional to the teaching that these children were otherwise receiving.

Our most notable result is that we have been able to demonstrate selective effects on these children’s reading skills. In line with the phonological linkage hypothesis, we have shown that an effective way of improving reading skills involves a joint approach that integrates the training of phonological skills with the teaching of reading. Spending an equivalent amount of time concentrating on either component in isolation (reading or phonology) is less effective. Although the individual teaching of reading received by the Reading Alone group did produce some gains, they were not as large as in the group given both reading and phonological training. This is an important, and not at all obvious, result. Generally the most effective way to teach a given skill is to teach it directly. Our children given the reading and phonology package actually received
less time being directly taught reading skills than did the Reading Alone group. The fact that they nevertheless made significantly more progress in reading is quite surprising and impressive.

According to our phonological linkage hypothesis, it is crucial that in the $R+P$ group explicit links were formed between reading activities and phonological knowledge. To this end the children in this group undertook linkage activities such as relating spellings to sounds using plastic letters and writing words while paying attention to letter-sound relationships. In addition, of course, the $R+P$ group received instruction in reading (comparable to the R group) and phonological training (comparable to the $P$ group). A skeptic might argue that the explicit linkage activities are not crucial to the success of the $R+P$ group and that instead children in this group might abstract the relationship between print and sound once they have some level of exposure to both phonological and reading exercises. This is certainly a possibility that our data cannot refute. We would, however, expect separate training in reading and phonological skills to be less effective than the explicit linkage given to the $R+P$ group. The study of Byrne and Fielding-Barnsley (1989), described in the introduction, supports this idea. They looked at young children's understanding of the alphabetic principle, the concept that particular phonemes in words are represented systematically by particular letters.

As we noted earlier, Byrne and Fielding-Barnsley found that such understanding was achieved only by children who could perform phonemic segmentation, understood phoneme identity, and had also been taught explicitly the critical phoneme-symbol relations (that S says /s/ and M says /m/, e.g.). This training of phoneme-symbol relations is an example of what we have termed a phonological linkage exercise in that it forces children to relate their awareness of phonemes to the process of reading words. Byrne and Fielding-Barnsley found that such training was necessary for their young children to come to understand the alphabetic principle. To test the critical role of our phonological linkage exercises more directly in the context of our own study would require a further study in which children received both reading and phonological instruction, but without explicit linkage exercises.

Leaving these details aside, our results certainly provide support for the view that phonological training alone is not a powerful way of improving children's reading skills. As we discussed earlier, Bradley and Bryant (1983) trained children in sound categorization and found that the gains in reading that resulted were not significantly greater than in a control group trained in semantic categorization. Similarly, Lundberg et al. (1988) found that training phonological skills in kindergarten children produced small effects on their later progress in learning to read. Our own results from older children with quite marked reading difficulties provide further evidence that improvements in phonological skills in isolation do not translate directly into improvements in reading skill. These findings therefore cast doubt on the simple theory that there is a direct causal path from phonological skills to reading skills. Our data support the more subtle position that adequate phonological skills may be necessary, but not sufficient, for learning to read effectively.

The phonological linkage hypothesis, and the support we have obtained for it, fits in well with a recent study concerned with the effectiveness of Clay's (1985) reading recovery program (Iversen & Tunmer, in press; Tunmer, in press). In Iversen and Tunmer's study a reading recovery program was compared with a modified program in which children also received systematic phonological training. The modified program was found to be more effective than the reading recovery program alone, just as our results would lead us to expect.

Intervention studies are notoriously difficult to conduct. We should like to stress a number of aspects of the design of the present study that help to strengthen the conclusions that can be drawn from it. One central point concerns the specificity of the effects obtained. The effects have been shown to be specific in two ways. First, we have shown that the improvements obtained are not completely general: there was not an equivalent pattern of improvement in arithmetic. This rules out nonspecific factors, such as motivation or teachers' expectations, as explanations of the results obtained. Second, we have shown that the beneficial effects on literacy development derived from an integrated approach to teaching reading and phonology are not purely mediated by changes in phonological skill. We produced larger effects on phonological skills in the Phonology Alone group without having any significant effect on literacy skills. This is not to say that improving phonological skills
in children with reading difficulties is unimportant. Our results do show quite clearly, however, that working on phonological skills in isolation is not an optimal method for improving literacy skills. This, of course, is exactly the pattern of results predicted by the phonological linkage hypothesis: teaching both phonological and reading skills and their interrelationship is far more effective than working on either in isolation.

Though our primary focus in this study has been on reading, we have also looked at spelling. Theoretically it has sometimes been argued that reading and spelling develop partially independently and that phonological strategies are more important for the development of spelling than the development of reading (Goswami & Bryant, 1990; Snowling & Hulme, 1991). If this were the case, we might expect slightly different effects of our interventions on progress in spelling than on progress in reading. In particular, if phonological skills are more intimately involved in learning to spell than in learning to read, it might be expected that phonological training alone would be of more benefit to spelling than reading. Our results obviously do not support this position. The P group did not make significant progress in learning to spell, while the R + P group made significant progress in spelling as they did in reading. Our results cannot refute the idea that learning to read and learning to spell are partially separate processes, but they do suggest that there is a strong association between the learning of these two skills. For children having difficulties in the early stages of learning to read and to spell it appears that an integrated package of reading and phonological training can be expected to improve both their reading and spelling skills, though gains in spelling appear rather smaller and less durable than those for reading.

Our main focus in this paper has been to demonstrate the effectiveness of one particular approach to teaching children with reading difficulties. One obvious question that arises, however, is the extent to which the effects obtained in our study will generalize across subjects. The children in our study were a representative group of poor readers corresponding, roughly, to the bottom 25% of reading skill. These 7-year-olds were struggling with the earliest stages of learning to read. It seems likely to us that the effects obtained would show wide generality. We know, for example, that children of high IQ with highly specific reading difficulties commonly experience underlying phonological difficulties (see Hulme & Snowling, 1991, for a review of this evidence). There seems every likelihood that such children with more specific reading difficulties would show an equivalent pattern of benefit from an integrated approach to teaching reading and phonological skills. It would, nevertheless, be useful to confirm this empirically. It is certainly possible, however, that a minority of children has such severe phonological difficulties that training in these skills will be relatively ineffective for them (see Hulme & Snowling, 1992; Snowling & Hulme, 1989; and Stackhouse & Snowling, 1992, for a discussion of some examples of such cases). This is an important issue for future intervention studies to address. Lovett, Warren-Chaplin, Ransby, and Borden (1990) failed to find a differential effect of letter-sound versus whole-word teaching methods with a group of dyslexic children. They did not, however, include any measures to improve these children's phonological skills of the sort used with the R + P group in the present study.

Although we believe our results will show considerable generality, it is possible, indeed likely, that individual differences among children will interact with the effects of teaching that they are given. We have data pertaining to this which are too complex to present in full. In a principal components factor analysis with VARIMAX rotation of all the 19 measures obtained at t1, we found six factors, the first three of which were Reading Ability, Phonological Ability, and Verbal Ability. We computed a composite measure of progress in reading between t1 and t2 based on scores from three of our reading measures (BAS, Neale accuracy, and Neale comprehension). The predictors of reading progress differed across the groups in an interesting way. For Group R the best predictor (r(29) = .72) was phonological ability at t1 (neither pre-intervention reading ability nor verbal ability were significant predictors: r's(29) = .14 and .20, respectively). For Group P, on the other hand, the best predictor (r(28) = .54) was reading ability at t1 (neither pre-intervention phonological ability nor verbal ability were significant predictors: r's(28) = .30 and -.02, respectively). For Group R + P the best predictor (r(30) = .54) was again phonological ability at t1 (neither pre-intervention reading ability nor verbal ability were significant predictors, r's(30) = .33 and .14, respectively). These results indicate that for children given only help with reading their pre-existing level of phonological ability is an
important determinant of their success in learning to read. On the other hand, for children only given help to develop their phonological skills, their pre-existing level of reading skill will exert a powerful effect on whether this intervention is effective in improving their reading skills. These results complement our phonological linkage hypothesis in indicating that phonological and reading skills need to be united if interventions to help poor readers are to be effective.

Another issue concerns generality across ages. Our children were on average 7½ years old and in the early stages of mastering reading and spelling skills. It is possible that training phonological skills will be more important for young children than older children. This is an important developmental issue for future studies to address.

We hope that the educational implications of our findings will be obvious. We have evidence from a controlled study showing the effectiveness of a structured teaching procedure that unites the teaching of phonological and reading skills. The amount of extra teaching received by the children in our study, although not at all trivial (40 30-min sessions over 20 weeks), is at a level that makes our intervention educationally realistic. The magnitude of the gains achieved also makes them educationally (as well as statistically!) significant. One slightly disappointing aspect of our results, however, is that the scores at the final follow-up (t3) suggest that the gains made by the R + P group, although still significant in reading accuracy and comprehension, tend to diminish. It may be that poor readers such as the children in our study require continuing support to recover fully from their early reading problems.

The results of our study also relate quite directly to a very large number of educational studies that have been concerned with the importance of phonics in the teaching of reading. In such studies, comparisons have typically been made between teaching methods that concentrate on reading for meaning and exclude phonics and those that make heavy use of phonics. It is obviously hard to exert rigorous control over irrelevant factors in such studies. However, the outcomes of a number of such studies (for reviews, see Adams, 1990; Chall, 1983) are impressively consistent in suggesting that phonic-based methods are more effective than meaning-based methods (though no one, we hope, would wish to belittle the importance of children learning to read for meaning). Our small-scale experimental study might be seen as providing evidence that complements these educational studies. In our study, poor readers in the R + P group certainly benefited from a remedial program that included many elements of phonics-based methods together with explicit training in phonological skills. In contrast, the R group, who received a highly structured approach that emphasized the use of context and meaning but omitted direct phonic teaching, was less successful.

One issue concerning the educational implications of our study is raised by the fact that all teaching was carried out on an individual basis. This is likely to maximize the effectiveness of any intervention but may also place constraints on how widely the methods could be used. Some of the procedures used in our intervention clearly need to be conducted on an individual basis (e.g., children reading and having their errors corrected by the teacher). However, some of the key elements of the R + P program, such as relating letters to sounds, making explicit links between the sounds of words and their spelling patterns, and certain phonological training exercises could be modified for use in small groups: this would be a useful extension of the present research.

Theoretically, the finding of support for the phonological linkage hypothesis raises a number of issues relevant to the normal development of reading skills and how best to facilitate them. Most obviously it raises the question of to what extent links between phonology and reading need to be made explicit in the teaching of reading. It also raises the question of to what extent children may differ in the ease with which they bring to bear the phonological skills that they possess in the task of learning to read and spell. As well as differing in the degree of phonological competence they possess, children may also differ in their ability and propensity to access this competence. This may represent another dimension of individual differences that contributes to differences in learning to read. If this is so, an important and parallel educational question is how children's access to underlying phonological skills can be facilitated in the teaching of reading.

References


Bruce, D. J. (1964). The analysis of word sounds by young children. *British Journal of Educational Psychology, 34,* 158–70.


This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.