Reading and Writing: An Interdisciplinary Journal **14:** 423–457, 2001.

The effect of phonics instruction on the reading comprehension of beginning readers

VINCENT CONNELLY[1], RHONA JOHNSTON[2] &
G. BRIAN THOMPSON[3]
[1]*Psychology, South Bank University, London, United Kingdom;* [2]*Psychology, University of St. Andrews, Fife, United Kingdom;* [3]*Education, Victoria University of Wellington, Wellington, New Zealand*

Abstract. This study investigated whether two groups of 6-year-old beginning readers taught to read by a phonics and by a "book experience" non-phonics approach would differ in reading comprehension as well as the processes of word recognition. The two groups were matched for word recognition but despite this, the phonics taught children had higher reading comprehension. Phonics taught children produced more contextually appropriate errors, and in both single word and text reading made more spoken attempts at reading unknown words. The non-phonics taught children had much faster reading reaction times to familiar words but they scored less in phoneme segmentation and nonword reading tasks. Compared with the non-phonics group, the phonics group spent more time in attempts at identifying unknown words and this included using contextual information, which apparently resulted in more rehearsal of the meaning of the story text and hence better reading comprehension performance.

Keywords: Comprehension, Context, Instruction, Phonics, Rate, Teaching

Introduction

The development of good reading comprehension depends on fluent word recognition skill (Adams 1990; Perfetti 1985; Stanovich 1980). There are a number of studies which have claimed that teaching containing phonics input advances beginning readers faster than teaching without phonics input. Phonics instruction is here defined as explicit teaching about individual letter-sound correspondences, their sequences (including spelling patterns), and the pronunciation of corresponding sounds. The phonics taught beginners in many of these studies were better at comprehending what they read than the non-phonics taught children. For example, Chall (1967) found that phonics teaching produced readers who had an advantage in word recognition and that by the end of second grade also had higher levels of comprehension and vocabulary than children taught by other methods such as 'look-say'. Bond and Dykstra (1967) in the co-operative research program presented a more

mixed picture but the general results showed that phonics was associated with faster development of word recognition and in some cases higher comprehension levels by the start of third grade. Adams (1990) also documents other studies which she claims support the conclusion that phonics teaching encourages faster word recognition and comprehension development compared to non-phonics teaching regimes.

Therefore, in those studies where the phonics taught children had higher reading comprehension they also had better word recognition skills than the non-phonics taught children. It could be argued that the beginners advantage in reading comprehension could be entirely due to the phonics taught children being better at word recognition than the non-phonics taught children. Word recognition skill has been shown to be the single best predictor of reading comprehension (Stanovich 1990; Vellutino 1991) and of particular relevance for this study, Curtis (1980) found reading comprehension skill to be highly related to word recognition in second grade.

On the other hand, the reading comprehension advantage could also be due to the different ways of processing adopted by beginners receiving different methods of reading instruction. There have been a number of studies which have shown that the method of reading instruction does affect children's beginning reading and how they process written information. For example, Seymour and Elder (1986) found that children taught by a "whole word" approach, which involved learning lists of words, could not read words outside their classroom learning set in their first year of reading. The children were very constrained by the limited set of words they were exposed to and had minimal letter sound knowledge at the end of their first year. In fact, they were so constrained in their word recognition that some of them had difficulty reading the taught classroom words on the different medium of a computer screen. Seymour and Elder (1986) called these children "logographic readers" after Frith (1985). Frith (1985) theorised that the first stage in reading development involved the direct learning of a small set of words through memorisation of distinct graphic features.

Wimmer and Hummer (1990) studied children in Austria learning to read using a systematic phonics approach. These children produced nonwords and mispronunciations based on the letter sound relationships they had been taught about in the classroom. A number of other studies looking at children taught different degrees of phonics have produced similar results (Barr, 1975; Seymour & Evans 1992; Stuart & Coltheart 1988) and show children reading very differently to the Seymour and Elder (1986) sample.

Studies which make direct comparisons with different teaching methods are more rare and even rarer are those who control for word recognition skill levels when making the comparisons. There have been a number of

comparisons of phonics taught readers in Scotland with New Zealand "Book experience" taught readers (A variant of language experience where reading for meaning is considered paramount) while matching for word recognition (Johnston & Thompson 1989; Thompson & Johnston 1993, in press). These studies reported similar results to those mentioned above where phonics taught children were slower readers but who made more use of letter sound information in order to generate pronunciations for written words even to the point where it became a disadvantage (as in pseudohomophone reading). However, no direct measurement of passage comprehension was taken in these studies.

Elder (1971) compared a group of Scottish phonics taught children who were equivalent in word recognition skill to 3rd Grade "look and say" taught children in the United States. The phonics taught children were a year younger but had received a years extra worth of schooling. It was found that despite being equal in word recognition the phonics taught children were a year behind in rate of reading (equivalent to U.S. Grade 2 norms). This was put down to the phonics taught children slowly attempting to generate a response using letter-sound correspondence. The phonics taught children produced more pronunciation errors and more nonword errors than the "look-say" children which would imply they were trying more to generate pronunciations of words. No direct measure of text comprehension was taken but error analysis was conducted on substitution errors. The "look and say" children produced more real word substitution errors and this was interpreted as being representative of a more contextually based approach to reading. However, it could also be interpreted as the "look say" children not attempting to obtain a reading response using generation procedures once recall has failed and so using real word substitution as the only available option. Elder found a trend towards the "look and say" children making more substitution errors that retained the words' original meaning. This led Elder to the conclusion that phonics teaching was detrimental to comprehension skill and that it "tends to divert children from the meaning of what they read" (Elder 1971, p. 228). However, since there was no direct measurement of comprehension of text and Elder did not examine pronunciation errors for meaning change (pronunciation errors constituted another 25% of errors in the phonics sample) we consider that Elder's conclusion may be premature.

Lesgold et al. (1985) compared cohorts of children taught by a phonics regime (including synthetic phonics, i.e. blending) and by a "global approach" consisting of a mixed basal and language experience curriculum from 1st to 3rd grade in the United States. Testing was based on when a child reached a level of the curriculum rather than time of year and so no direct comparison of comprehension was made due to the cohorts receiving

comprehension tests at different times of the school year. In single word reading, as in the Elder study, the phonics taught children were slower readers during Grades 1 and 2. The phonics taught children then caught up in reading speed with the global approach children in Grade 3. The phonics taught children also produced more nonword errors. There was no analysis reported of meaning related errors. Lesgold et al. (1985) did include a semantic judgement task that was directly comparable across instructional cohorts. This task showed that the phonics cohort was initially as accurate as the global method children if a bit slower to respond. By Grade 3 the phonics taught children had become as fast as the global method children but were now less accurate. There was no obvious explanation for this result.

In summary, the aforementioned studies have shown that different reading instruction influences young children's reading patterns even when they are matched for level of word recognition attainment. It seems that phonics taught children in the early years of instruction use their training to attempt to generate pronunciations for words. This makes them slower readers who are liable to produce mispronunciations and nonwords as errors. However, as illustrated above, it is not clear how comprehension may develop or be affected by differing instructional methods. It is necessary to carry out studies where non-phonics and phonics groups are equated for word reading attainment in order to examine how differences in their reading approaches may impact on factors such as comprehension. From the evidence quoted above, the best time to achieve a word reading match is during the second year of reading instruction.

The present study compares two groups of young readers who were matched on word recognition level, but who differed widely in the type of reading instruction they have received. One group had been trained very systematically in phonics and had been encouraged to sound out words using their knowledge of letter-sound correspondences. The other group had received "book experience" instruction; this is a method of instruction, in which phonics or letter sound correspondences are not explicitly taught, but the child is encouraged to use context to aid word recognition. For matched word recognition levels, does teaching children to use a book experience approach in reading, where the use of context is emphasised for word recognition, lead to better reading comprehension than a teaching method where children are taught to use letter-sound correspondences to decode words? That is to say, over and above any advantage to comprehension from better word recognition, is there something inherent in the different processes acquired by children with and without phonics instruction which produces comprehension differences?

Method

Participants

Four schools were used for the study. Two of the schools were situated in Tayside, Scotland and taught an intensive phonics programme while the two schools in the Wellington district of New Zealand taught beginning reading according to a non-phonics "book experience" programme prevalent throughout that country. They were matched as closely as possible on socio-economic status (SES) from within two small cities of comparable size. The first author selected the Scottish schools which were in an area defined by a majority of lower middle SES parents (about 60–70%) but which also included middle and higher middle SES parents (10–20%) and lower SES parents (10–20%). The first author had lived and worked in this area for many years and was very familiar with the mixed SES of the area. The SES was confirmed by talking with the headteachers and the class teachers about the SES composition of the classes studied.

Scotland and New Zealand have many historical, cultural and social connections and similarities of urban culture. English is the majority language in both countries. Comparison of the methods of teaching of reading in the two countries allows us to ensure that no classes were taught by teachers who had received training in another method of teaching reading as is common in some countries such as England (where teachers are trained on a more regional basis rather than a national basis) and in the United States. It was therefore quite easy for the third author, who has worked in New Zealand education for many years to identify two schools which were of equivalent SES status in New Zealand after discussions with the first author. Again the SES status was confirmed by talking to the headteacher and class teachers of the schools in New Zealand. The first author worked in the New Zealand classes as well as the Scottish classes and so was able to personally confirm there was no obvious SES differences.

Both Scotland and New Zealand have professional career teachers who follow a similar career structure. All the teachers co-operating in this study were graduates and all had at least 5 years teaching experience with the age groups involved. One school in each country had the deputy head teacher of the school as one of the teachers involved in the study. The staff/student ratios in each country are very similar and the mean ratio for the classes used in Scotland was 1/26 and in New Zealand 1/24.

Schools in both countries also encouraged parents to come into the classroom to act as assistants during lessons. In practice this was quite rare and in Scotland about one or two parents for one morning per week was average. In New Zealand there was slightly more participation with

one consistent parent two mornings a week and one parent one occasional afternoon a week in the classes observed.

Reading instruction

Phonics regime

Educational approach. The Scottish children received a heavy diet of formal phonics lessons reinforced by a phonics based reading scheme. The headteachers and teachers of the Scottish schools were very keen to point out that the rapid build up of phonics knowledge was of utmost priority to them in the initial stages of reading instruction. They believed that only a solid understanding of phonics rules and the alphabetic principle would serve as basis for future independent reading. This belief has been current in Scottish schools for many years. Elder (1971) mentions it in reference to his data which was collected in 1962.

It is important to note that phonics teaching while traditionally strong and very prevalent in Scotland is not necessarily universal. However, any non-phonics implementations are generally done on a very local basis under the auspices of the headteacher. Chall (1967) reported that teachers who had taught in non-phonics regimes could carry old practices into phonics regimes. The first author questioned all the Scottish teachers and none of them had previously taught in a non-phonics regime and they had always used phonics approaches.

Considerable autonomy is given to the Scottish schools to plan their reading curriculum with general guidelines and recommendations issued by the respective education authority that the schools work within. Reading schemes are used but the reading curriculum is not strictly tied to any one reading scheme and lessons are not planned rigidly from the materials and lesson plans provided by reading scheme publishers.

Phonics teaching. Most direct phonics teaching consisted of stand alone lessons taught to the whole class which were distinctly separate from the reading scheme work. A typical lesson would begin by the teacher introducing a sound to the class verbally while pointing to the appropriate letter or letters on a chalkboard. The class would look at the chalkboard and verbally repeat the sound as a whole a number of times. If the sound constituted a "phonics rule" such as the vowel blend "oo" as in "good" and "food", then the children would repeat the rule out loud, in this case, "two 'o's make 'oo'." The children would sometimes be asked individually to read out loud the sound to the rest of the class. The teacher would then read out a list of words written on the board in which the sound or rule was present. This was usually

about 10–20 words. For example, for the sound ‘ch’, the following list was observed; chi, cho, che, cha, chu, chat, charm, chill, chop, chest, check, chin, ditch, hutch, patch, lunch, bunch (note the use of nonword stimuli). The letters in the words would be colour coded to highlight the sound being learnt. The children would read out loud the words emphasising the separate sounds they knew. This process of “sounding out” the words would be continued for up to 10 minutes as a class, led by the teacher. The children would then be asked to contribute more words containing the critical sound. If correct, the class would then repeat these as well. A selection of individual children in the class group would be asked to sound out loud some of the words on the board to the rest of the class. A lesson such as this would last in total about 15–20 minutes.

Generally the lesson was once a day but was observed occasionally twice in a day. How many times the children would receive the lesson was dependent on how complex the sound was considered by the teacher and how well the children performed on the follow up phonics worksheets. The same lesson for simple sounds such as single consonant sounds were repeated two or three times during a week. More complex sound lessons could be repeated for five or six times.

There were worksheets which were completed by the children after the whole class lesson. These were a mix of commercially produced phonics material, material from the reading scheme, GINN 360 or were prepared by the teacher. The worksheets engaged the children in written exercises designed to reinforce the whole class lesson. They varied in style widely and used a variety of methods to assess and reinforce the phonics lessons. For example, the child may have been asked to circle all the words that contained the sound within a page of words. Other worksheets might require the child to fill in missing words in sentences with an appropriate word that contained the sound being worked on. Still others might require the child to produce as many words with the sound as possible or the child would be asked to choose the correct sound from a list of written choices to fill the gaps in portions of words listed on the page. The worksheets encouraged the child to analyse words for the target sounds and also required them to synthesise sounds to make up words.

The letter-sound correspondences covered by Primary 2, the second year of instruction, included all single consonant letter sounds, consonant blends (such as ‘sh’, ‘ch’ and ‘ng’), simple vowel sounds (such as ‘o’ in dog and ‘a’ in ‘cat’) and the more complex vowel blends such as ‘ae’ and ‘oo’ as well as ‘magic e’ (such as in ‘gate’). At the time of testing more complex phonics rules relating to sounds like ‘igh’, ‘silent b’ (as in ‘lamb’) and ‘soft c’ (as in ‘pencil’) were being taught.

Text reading. Text reading was done using publishers reading scheme books. The primary reading scheme used was GINN 360 (GINN 1988a, b) which has a phonics emphasis. The scheme claims to emphasise comprehension and vocabulary development. The scheme is presented in a number of levels from 1–13. Each level has a set of books associated with it. The children studied were undergoing a transition from level 4 to level 5. Level 4 and 5 each had six main books ('readers') associated with it. Level 4 also had 12 other associated books ('little readers') to be used as supplementary texts to reinforce the vocabulary encountered in the main 'readers'. Level 5 had 6 supplementary books (titled 'magic circle books'). The books were a mix of genres including fiction, poetry and non-fiction. Beck and Block (1979) reviewed GINN 720, a revised scheme due to be introduced in the United States in 1980 to replace GINN 360 which was published in 1969. The British version of GINN 360 was published in 1979 and incorporated almost all of the changes the US GINN 720 implemented. Beck and Block (1979) concluded that the GINN stories were interesting, varied and of high quality and provided a realistic picture of a pluralistic society. The amount of time the child spent with each book would vary according to its length and complexity. Level 4 books were generally used for no more than a week while Level 5 books could be used for up to 2 weeks as they were much longer and included several narratives.

GINN 360 introduced a defined set of words at each level. These words were designed to fit in with the programme of letter-sound correspondence instruction recommended for each level of the scheme. The programme of letter-sound correspondence instruction in GINN 360 was slower than that experienced by the children in their phonics lessons. The scheme also introduced high frequency words that contained exceptions to the learned letter sound correspondences (e.g. words such as 'you'). Repetition of words was prominent in the stories included in the books. For example in the Level 4, Book 5 story "Little Monkey", the word 'tree' is repeated 7 times and the word 'friend' 10 times.

Worksheets to supplement the reading books were also used in the classroom. These emphasised letter-sound correspondences and single word repetition as described above. Most of the worksheets could be used alone without having directly read the associated books. This allowed the teachers to use the letter-sound worksheets at their pace of phonics teaching and not the reading schemes pace.

Comprehension questions were recommended by the publishers for teachers after children had read a story. These questions were generally 'wh' questions ("What is the lion doing?", "Where do the monkeys live?"). A few

'challenge' questions were used to supplement the general questions (e.g. "Tell me how you know that?")

Comprehension worksheets were also used to assess understanding in the children. These worksheets included tasks such as matching appropriate sentences with pictures or providing written answers to questions about a picture. The worksheets were closely tied to the reading scheme books due to the knowledge the child was assumed to have derived from the reading book story and so were used at the same pace as the reading scheme. The reading scheme recommends that teachers encourage the child to use context to aid prediction when reading the stories in addition to using letter-sound correspondence. This approach was not observed in the Scottish teachers. They always encouraged the child to sound out any unknown words.

Classroom practice. A typical school day for the Scottish children is shown in Figure 1. A reading period in the Scottish schools would usually begin with a phonics lesson as described above. Once this was complete the class would split up into ability based reading groups of 5–6 children. The teacher would brief the children on what activities they were expected to do and then ask one of the ability groups to gather round the teacher with their current reading scheme book. The group would then read the text out loud to the teacher, each child reading one or two pages. The teacher would ask comprehension questions about the text along the reading scheme recommended lines. Whenever a child did not know a word they were encouraged to generate a response using the letter-sound correspondences they had been taught. Emphasis was always at the letter-sound and word level in these text reading sessions. The teacher would point out words that included letter-sound correspondences they had covered in phonics lessons. Each group would read to the teacher for about 10–15 minutes. The group would then go back to their seats and continue with the other allotted reading tasks while another group began reading with the teacher. Children's reading of the GINN 360 texts was listened to at least twice a week by the teaching staff. In the authors opinion, the group reading sessions seemed more about evaluating the children's reading level and competence in decoding rather than explicit teaching about text reading and comprehension.

Those groups not reading with the teacher would be carrying out a number of different activities. These would include phonics worksheets or activities as described above, comprehension worksheets related to the reading scheme stories or silent reading of reading scheme books or more rarely other commercial books chosen from the small class library. Generally each group would have a similar pattern of work to complete during the allocated reading period. This would be two phonics worksheets or activities first, followed

by silent reading, interspersed with reading to the teacher as a group. A comprehension activity might replace one of the phonics activities two or three times a week. One or two groups every day would also filter a pair of children, for up to 15–20 minutes at a time, to the class computer. The pair of children would use commercially available literacy software on the computer. The majority of the software consisted of phonics programmes but there were a few comprehension and writing programmes available.

The classroom teacher would call the class together at the end of the time allocated to reading to either read a story to them or discuss a phonics point that needed reinforcing. Reading and literacy activities averaged about 90–95 minutes of the school day in total.

Written language. Writing took up a large portion of the literacy related activities. The reading worksheets depended on the child reading but generally involved a written response. Time was allocated most days to story writing, usually near the start of the day. The children had story books in which to write. The children were encouraged to write their own stories but if the child wished to write a complex sentence then the teacher would write it out and the child would copy it into their story book. Children were encouraged to write more stories or complete what they had started when they had completed their reading activity work.

Formal spelling instruction did not begin until the second year of instruction. Until then, invented spelling was tolerated. Formal spelling instruction involved the children learning lists of words and being tested on them by the teacher. The lists of words presented for spelling instruction generally did conform to the phonics instruction the children had received but a proportion of the words taught were high frequency exception words. In their written work, invented spelling was tolerated for words not yet taught but they were corrected by the teacher. The children were encouraged to get the teacher to write correct spellings for them in their own personal spelling dictionaries. They then used these dictionaries when writing.

Home environment. No observations were made of the children's home environment but the children were encouraged to take books home most evenings to read with their parents. The teachers indicated that most parents did read these books with their children. No written guidelines were issued to parents from the schools about the schools stance on reading instruction that was seen by the authors. However, both schools held regular evening meetings about various aspects of the curriculum and teaching throughout the year which any parent was welcome to attend.

Phonics Classroom		Non-phonics Classroom	
Approximate Time	Typical Activity	Approximate Time	Typical Activity
08.55-09.20	Registration and Religious Education or story writing	09.00-09.15	Registration and Religious Education
09.15-09.30	Spelling or story writing	09.15-10.00	Maths Activities
09.30-10.30	Maths Activities	10.00-10.30	Story Writing
10.30-10.45	Break	10.30-10.50	Break
10.45-12.15	Reading Activities	10.50-12.10	Reading Activities
12.15-13.25	Lunch	12.10-13.20	Lunch
13.25-14.30	Arts Activities (Also completion of reading and maths work from morning)	13.20-13.45	Arts Activities
14.30-14.40	Break	13.45-14.15	"Choosing" Activities including writing poems, listening corner and 10 minutes whole class story reading by teacher
14.40-14.55	Arts Activities (Also completion of reading and maths work from morning)	14.15-1500	Physical Education
14.55-15.05	Whole class story reading by teacher		

Figure 1. Typical school day structure based on classroom observation.

Book experience

Educational Approach. Reading for meaning is one of the primary goals of the New Zealand "book experience" approach and it is recognised that there is a distinct "New Zealand style for teaching reading in the early years" (New Zealand Department of Education, 1985 p. 7). The philosophy of the New Zealand teaching style is compatible with the psycholinguistic reading theories of authors such as Smith (1978) but there has been an emphasis on obtaining meaning from print in New Zealand for more than 40 years (New Zealand Department of Education 1985; Thompson 1993). From 1962 onwards in New Zealand there has been a shift from emphasis on words in reading teaching to an emphasis on the story and the book. This has become more prevalent in the last twenty years and there has been increasing concern that children are able to predict reading responses from story and sentence contexts. The New Zealand classes we studied contained teachers who were instructed in New Zealand methods at New Zealand teacher training institutions and who agreed they had always used the New Zealand style for teaching reading.

New Zealand schools use a common core reading series, Ready to Read, which was published and trialled by the New Zealand Department of Education (1985) in the early 1980's. The series is finely graded and has a sequence of nine levels covering the initial two to three years of instruction. The selection of scripts was judged on the appeal of the story to young readers and the relevance and significance of the content. The second year level texts are published as single titles or small collections of texts. The texts cover a

range of genre: fiction, nonfiction, poems, fables and plays. The books have some vocabulary in common, particularly at the lower levels, but there is very little control of the rate of introduction of new words to the children. In addition to the Ready to Read texts many commercial print books graded for the appropriate reading levels are available in the classroom for use.

Text reading. The children in New Zealand and in the schools studied encountered a single text in the three different book experience methods of shared book reading, guided reading and independent reading. These methods ran concurrently throughout the teaching programme. Books were given their own grade level for each method of teaching use. Some books received recommendations for some methods but not others.

Shared reading

Shared reading involved the teacher reading a text with the whole class or a group of children. The teacher typically used a "Big Book" of enlarged size which the children could watch as the teacher pointed to the text. Prior to reading the text the book was introduced to the class by the teacher discussing the general story topic but without revealing the actual story line. The children were encouraged to predict what the text would be about and what would happen in the text from the title of the book or story. The teacher then read the text out loud to the children and paused to discuss predictions about events and words or conventions of print where it was considered appropriate. The main purpose of shared reading was for the child to enjoy the story in order to encourage an interest in reading. It also provided the teacher with the opportunity to introduce unfamiliar words and concepts to the children in what was considered a non-threatening manner.

Guided reading

Guided reading with a child or small group of children generally involved the children following a story with their own copy of the text and the teacher actively working with the children in predicting words and events in the story. The children were sometimes encouraged to use letter name cues to assist in word recognition. The children were encouraged to read silently in these sessions. Silent reading was followed by the teacher asking comprehension questions to gauge how well the children had attended to the meaning of the text. Guided reading was teacher led and was not the teacher passively listening to children's reading. Guided reading texts were generally used for about a week per child before another was selected by the teacher.

Independent reading

The child always had a book available for independent reading. This was an easier grade level book that the child selected from a range of titles and which had been experienced in shared and guided reading practice previously. The children engaged in reading this book daily in the class when asked to do some independent reading by the teacher. It was also noted that a number of children would read their book when "choosing" their own activities in the classroom. The children were encouraged to read their book silently. The aim of independent reading was for the child to practice successful reading without external assistance so providing an enjoyable reading experience. In the schools that were observed by the authors all the children took independent reading books home every day.

Phonics philosophy. Phonics teaching has not been common in New Zealand schools for 30 years or more. There is some implicit phonics recommended by the education department which is based on the teaching of letter names. When a child is struggling with a word and after they have attempted prediction of that word from context then it is recommended that the child use the initial letter name of the word to try to work out the word. Children receiving the New Zealand style of teaching reading do not overtly "sound out" or "blend" segments of words.

Classroom practice. The first author informally observed teaching in the New Zealand classes twice a week for the three months of the testing period. At no time during this period was any kind of explicit phonics related activity, apart from encouragement to use letter names at the beginning and end of words, ever observed. No "sounding out" or "blending" was modelled by the teaching staff or parent helpers. Most teachers did, however, encourage their children to use initial letter names to infer word recognition in conjunction with contextual cues rather than after all contextual cues had been exhausted. During guided reading, the teachers were observed directing children's attention to listening to initial common sounds associated with print words, e.g. big, baby. This was the only major deviation from the New Zealand Department of Education (1985) guidelines observed during the study period and constitutes minor implicit phonics.

A typical school day for the New Zealand children is illustrated in Figure 1. The reading practice would usually begin by the teacher engaging in shared reading for 10 minutes with the whole class used an enlarged book. The class would then split down into reading ability groups of 4–6 children each. Then, for example, with five groups of children, one group would engage in guided reading with the teacher, two groups would engage

in silent independent reading, and two groups would complete worksheets relating to the texts they had been reading. These worksheets were comprised of comprehension and meaning related questions. After fifteen minutes the groups would change around. The teacher would take another group for guided reading or might work with one or two individuals or help a group with the worksheets. This routine would continue for some time until the teacher would call the class together to discuss points arising from shared or guided reading or allow the children to describe what they had produced in their worksheets or other tasks they had been set.

The teacher would spend a proportion of time per week listening to individuals in order to construct a "Running Record" of a child's reading progress. The teacher obtained an error rate per running words of text and interpreted the quality of children's responses with special reference to comprehension of the story-line. The result was then used to match the child with appropriate texts. A running text appropriate accuracy rate for independent reading is 95% or greater and guided reading should have accuracy rates between 90–95% (New Zealand Department of Education 1985). A "Running Record" was completed for each child once a month and more often for poorer readers.

The detailed "Running Record" allows for great individualisation of instruction in New Zealand and is seen as an important part of the programme. Since children in New Zealand start school on their fifth birthday, there is a constant demand on teachers to fine tune instruction for the individuals who arrive randomly throughout the first year of school. After one year of instruction, at age six, those children in the slowest 20% of readers are programmed to receive intervention in the form of the Reading Recovery programme (Clay 1979). This involves daily withdrawal from the classroom for 30–40 minutes to work individually with a Reading Recovery trained tutor. Once the child has been judged to have reached a level of reading performance at or above the class average they are returned to the classroom. This usually takes from 12–20 weeks.

In summary, while the complete teaching approach in New Zealand can be regarded as compatible with 'whole language' approaches, the finely graded texts, detailed record keeping and "Reading Recovery" back up are not always present in other implementations of 'whole language'. The teaching approach is not similar to that described by Seymour and Elder (1986) and it was not expected that the children under instruction would be logographic readers as Seymour and Elder describe their sample. The classes were tightly controlled by the teachers who were firmly in control of the pace and content of lessons and pupils had a large amount of teacher contact time in their reading lessons. The New Zealand teachers were very keen on

keeping children's attention on their allotted task through constant observation of the children in the classroom. This was confirmed in our observations of the New Zealand classrooms. Glynn et al. (1989) also report that New Zealand teachers ensure that children are "on task" for a large percentage of the reading lesson and that New Zealand schools are consistent in their teaching approaches.

Reading and literacy activities averaged about 90-100 minutes of the school day in total for the book experience regime from observed classroom lessons.

Written language. Reading takes up the largest part of literacy related activity but there is still a heavy emphasis on the child using written language. Children were encouraged to write their own stories, messages and poems. Teachers encouraged the children to use invented spelling in the first year but in our sample of second year pupils teachers had begun to encourage the child to use dictionaries to spell words correctly. It is thought by teachers that there will be transfer of letter sound knowledge from writing to reading when children learn to spell in New Zealand. Thompson (1993) reports evidence which challenges this belief.

Home environment. No observations of the home environment were made but the New Zealand children were encouraged to take books home to share with their parents. Parents received a leaflet from the school informing them of the Education Departments policy on reading, how children learn to read and how they may assist their child in learning to read.

Summary. As Chall (1967) and a reviewer of this paper pointed out, the difference between teaching methods in practice should not be overestimated or overdichotomised. We acknowledge this and realise that the main difference between both programmes described above is one of when and to what degree letter sound correspondences are introduced and with what emphasis. Both methods of teaching reading have the aim of producing independent readers who enjoy reading and who can continue to make further progress in language arts with a solid skill foundation in reading behind them. The difference is therefore more of a technical difference rather than a fundamental one. However, we agree with Perfetti (1985) who states that a technical difference in methods "should not be equated with trivial" (Perfetti 1985, p. 252).

Matching of groups: A matching procedure was used so that differences were not due to one group having better word recognition. The children were also matched on aural vocabulary knowledge which correlates with reading comprehension, and digit span, a measure of short-term memory.

Time at school and chronological age were also matched. These were the only measures collected from the whole class samples.

Pool of phonics taught children. Pupils were from two classes in their second year of schooling (ages 6 years to 7 years approximately). This gave a total potential number of 52 pupils from Primary 2. Forty five parents expressed their wish to have their child included in the testing. These pupils had a mean age of 6 years and 4 months and had been at school for an average of 14.8 months. They had a mean reading age of 6 years and 11 months on the British Abilities Scales (BAS) Word Reading Test (Elliott, Murray & Pearson 1977), a test of the oral reading of individual words. They had a mean British Picture Vocabulary Scale-Short Form (BPVS) (Dunn & Dunn, 1982) standard score of 96.7 and an average standard score of 117.7 on the BAS digit recall test.

Pool of non-phonics taught children. Pupils from the Junior Two class (ages 6 to 7 years approximately), were studied. This gave a total potential participant number of 48 pupils of suitable age.

Those for whom English was a second language (5) were removed from the sample. This left a total of 43 pupils. They had a mean age of six years and 5 months and had been at school for an average of 14.7 months. They had a mean reading age of 6 years and 10 months on the BAS word reading test. They had a mean BPVS standard score of 95.4 and a mean standard score of 118.07 on the BAS digit recall test.

Selection of final samples. The aim of selection was to have two groups of children who were equivalent in word recognition, vocabulary development, digit span, age and time at school. Since the means of the whole classes in these measures were quite close it was thought this would be easily achieved. However, on examination of the distributions of the whole class samples for word recognition, a distinct non-normal pattern was found. There were clusters of children at the bottom and top of the distribution in each class. Lesgold et al. (1985) also found high and low ability clusters in their two instructional cohorts. Their clusters comprised about 40% of the results in total for each cohort. Lesgold et al. (1985) found significant differences on all measures including rate, accuracy and comprehension between the samples of high, medium and low ability readers found in their study.

Furthermore, in New Zealand the children at the bottom end of the distribution were receiving or liable to receive "Reading Recovery" and so would not be eligible for a comparative study with the Scottish children (some of whom had been receiving individual help also). It was considered that the cluster of children at the top end of the distribution in each class was likely to be children who may have begun reading before school or who

Table 1. Means of matching variables for groups (standard deviations in parentheses)

	Phonics group	Non-phonics group
Variable	n = 22	n = 26
Chronological age (months)	76.1 (2.01)	77.4 (2.9)
Time attended school (months)	14.8 (0.3)	14.9 (2.4)
BAS 1 word reading age (months)	81.4 (4.1)	79.9 (4.8)
BAS 2 word reading age (months)	84.7 (5.0)	83.7 (5.6)
BPVS standard score	98.0 (11.5)	95.2 (6.1)
BAS recall of digits standard score	116.1 (20.8)	118.2 (12.9)

had been receiving intensive parental help (after Lesgold et al. 1985). Therefore, it was decided to match the largest group of Scottish and New Zealand children possible on all the factors without including the cluster of individuals at the top and bottom of the distributions. In practice this meant that individuals from the whole class samples who were within one and a half standard deviations of the BAS Word Reading mean were included in the final experimental sample.

Six phonics taught children and eight non-phonics taught children were eliminated due to being at the bottom of the distribution. Twelve phonics taught children and nine non-phonics taught children were eliminated due to being clustered at the top end of the distribution. Five phonics taught children were also removed as they were too young to compare with the New Zealand children. This was a consequence of New Zealand children starting school on their fifth birthday exactly and some Scottish children starting school before age five.

This led to the selection of two groups of children whose mean scores and ages are given in Table 1. They will be referred to as the phonics and the non-phonics groups. The groups were retested on the BAS Word Reading Test four months after the first test (hereafter referred to as BAS 1 and BAS 2) in order to show that the groups did not regress to different means either side of the period when the Neale Analysis of Reading Ability (Neale 1989) was administered.

The phonics group consisted of 9 females and 13 males while the non-phonics group consisted of 10 males and 16 females. All of the analyses run in the study were investigated for significant gender bias in the results. None was found in any of the analyses.

To evaluate the closeness of match of the groups, analyses of variance (ANOVA) were carried out to test for any differences between the groups. A

repeated measures ANOVA on the word reading test scores showed no main effect of group, [$F(1, 46) = 1.03$, $p > 0.05$] but the main effect of reading test (BAS1 & BAS2) was significant [$F(1, 46) = 86.2$, $p < 0.01$]. The interaction was not significant ($F < 1$). Therefore, there were no differences in word recognition ability between the two groups over a four month period and the scores of the samples increased significantly between BAS 1 and BAS 2 by an equivalent amount.

One way ANOVAS were carried out on potential differences in chronological age, time at school, British Picture Vocabulary Scales (Dunn & Dunn 1982) and BAS Recall of Digits (Elliot et al. 1977) scores. No significant differences were found between the groups for any of these factors ($F < 1$). Therefore, the two groups would seem to be closely matched in terms of word recognition, chronological age, aural vocabulary knowledge, recall of digits and time at school.

It is acknowledged that this matching procedure necessarily produces a sample that is not random, is artificial and that is not whole class. It was considered that the selection was worthwhile in order to investigate the experimental questions we had posed without confounding variables intervening (primarily from different word recognition and vocabulary skill levels).

Materials and procedures

The children were tested on reading comprehension using the Neale Analysis of Reading Ability – Revised (Neale 1989). Measures of word reading accuracy and reading rate are also included in this test. A number of other tasks were also given to the children to investigate reading processes (phonological awareness, regular word and exception word reading, reading reaction time for familiar words and nonword reading). All administration was carried out for both samples by the same researcher.

Neale analysis of reading ability. The Neale Analysis of Reading Ability-Revised (Neale 1989) was used to test the reading of continuous text. It is a widely used test in both Scotland and New Zealand and was familiar to all the teachers in the schools. Form 1 of the test was used for all the children. Each child received a practice narrative at the start of the test to familiarise them with test procedures.

The test itself comprised a series of six short graded narratives designed to test the reading rate, word reading accuracy and comprehension of children's oral reading between the ages of 6 and 12 years. The narratives were presented in a book and each narrative had "a central theme, action and resolution" (Neale 1989, p. 7). A picture accompanied each narrative. The

child's reading errors (including failures to attempt) were corrected by the tester as reading proceeded.

There were comprehension questions for each passage which were answered orally by the child after reading each narrative. The comprehension questions were designed to allow use of all possible cues including contextual and picture related. They tested the "immediate recall of the main idea of the narrative, the sequence of events, other details and some limited inference" (Neale 1989, p. 7).

The Neale Analysis was designed to be as natural a reading experience as possible using a book with pictures and interesting stories. The New Zealand children were quite used to an adult recording error data and timing reading while they were reading due to the many "Running Records" the teaching staff had completed with them.

Yopp Singer Phoneme Segmentation Test. This is a phonological awareness test (Yopp 1988) which measures the ability to articulate the sounds of a word separately in the correct order. The child was asked to say all the sounds that were in the aurally presented test word. Three examples were worked through with the child. There were 22 test items.

Reading words of regular and exception letter-sound correspondences. The children read 10 regular and 10 exception words taken from Holligan and Johnston (1991) and classified according to Venezky (1973). The sets of words were matched for word frequency according to the grade 3 norms of Carroll, Davies & Richman (1971) (Regular words, mean frequency 165.2, exception words, mean frequency, 239.6). The children read out aloud the words presented to them. The words are listed in Appendix A.

Reading reaction times for familiar words. A set of 60 words were first selected from the GINN 360 Level 1 and 2 Teachers' Resource Book (1988) and Thompson's (1982) study of the frequency of words present in reading books in New Zealand schools. It was established whether the words had been encountered in the day to day reading in class. On this basis, 40 words were chosen that would have been encountered by the children in both the phonics and non-phonics classes. The Grade 3 norms of Carroll et al., (1971) gave the mean frequency of the words as 5355 (SD = 2618). The words are listed in Appendix B.

An Amstrad Portable PC and a Microvitec CUB monitor (12 inch display) were used to present the stimuli. The PC was interfaced with a verbal response unit via an analogue sensitivity controller which allowed adjustment to accommodate the normal speaking voice level of individual children.

The same equipment was used in both Scotland and New Zealand and was calibrated and operated in the same way.

The stimuli letters were designed using software on a matrix of 20 (x) by 30 (y) pixels. All letters were in lower case and based on those typically found in early reading materials. An asterisk character was used as a fixation point. Reaction times were measured in milliseconds. The sessions were also recorded on audio cassette tape for analysis of accuracy of the responses.

The 40 word stimuli were randomised by the software for presentation. There were four practice trials. The children were told that they were going to see some words on the screen and that they had to say the word out loud as quickly but as carefully as they could.

Nonword reading. One letter was changed in 24 of the words used for the naming speed task to make it a nonword that would bear an orthographic resemblance to a real word familiar to the children. All the nonwords were clearly printed in lower case in the middle of white index cards. The items are listed in Appendix C.

The children were asked to read out loud the 24 words presented to them. Three practice items were generated in order to familiarise the children with the task. It was clearly explained that the nonwords were "made up words". The set of cards was shuffled for each child.

Results

Neale analysis of reading ability

Normative reading ages were obtained from the three measures of the Neale (see Table 2 which also details the mean age at the time of testing for the two groups). Raw scores from the test were also examined and exactly the same pattern of results was found for the raw scores as for the normative reading ages reported below.

Word accuracy measure. A one way ANOVA was performed on the word accuracy data for reading continuous text in this test. No significant difference between the phonics and non-phonics groups was found [$F(1, 46) = 0.211$, $p = 0.647$]. Since the groups were equated on word recognition in the BAS test this result is not unexpected. However, the reading age shown by the non-phonics taught children for word accuracy was significantly below their chronological age at the time of the test [$t(26) = 3.44$, $p = 0.002$]. This was also true for the phonics group [$t(22) = 2.82$, $p = 0.01$].

Table 2. Mean age at testing, reading ages in months for the Neale analysis of reading ability and reading rate in mean words per minute (Neale 1989) (standard deviations in parentheses)

	Phonics group	Non-phonics group
Measure	n = 22	n = 26
Age in months at Testing	78.8 (2.1)	79.8 (2.9)
Word Reading Accuracy Age	74.1 (6.8)	73.5 (8.2)
Comprehension Age	78.4 (6.9)	73.9 (8.0)
Reading rate Age	68.1 (9.9)	83.2 (17.8)
Reading Rate in mean words per minute	30.6 (10.5)	45.9 (17.0)

This may be a function of the Neale Reading Ages for Accuracy scoring almost a year behind normative ages on the BAS word recognition test and may not reflect accurate age appropriate ability.

Comprehension. Given that the phonics taught children scored slightly higher (but not significantly) in the word accuracy measure of the Neale test than the non-phonics taught children and there was a similar tendency for aural vocabulary (BPVS scores) an analysis of covariance was performed on the comprehension scores taking word accuracy and BPVS scores into account as covariates. A significant difference in comprehension remained after accuracy and vocabulary had been accounted for [$F(1, 44) = 5.63$, $p = 0.02$] in favour of the phonics group. The difference, analysed without the covariates was also significant [$F(1, 46) = 6.45$, $p = 0.01$]. It was also found that the non-phonics taught children had a reading age for comprehension that was significantly lower than their chronological age [$t(26) = 3.33$, $p = 0.003$] while the phonics taught children showed no significant difference between chronological age and reading age [$t(22) = 0.31$, $p = 0.763$].

Reading rate. A significant difference between the phonics and non-phonics groups was found [$F(1, 45) = 11.8$, $p = 0.001$] with the non-phonics taught children being faster readers. The phonics taught children were significantly lower in reading age for reading rate than their chronological age [$t(22) = 5.03$, $p < 0.000$] while the non-phonics taught children were not [$t(25) = 0.94$, $p = 0.355$].

Correlations between measures. As expected there was a significant positive correlation between comprehension and word accuracy in both groups (see

Table 3. Mean results for other reading and reading related measures (standard deviations in parentheses)

	Phonics group	Non-phonics group
Measure	n = 22	n = 26
Nonword reading% Accuracy	64.4 (17.8)	38.8 (19.7)
Yopp-Singer (1988) Phoneme Segmentation, number of items Correct	16.4 (3.4)	13.7 (3.8)
Regular Word Reading% Accuracy	67.3 (25.5)	52.4 (25.9)
Exception Word Reading% Accuracy	32.5 (19.4)	41.0 (24.0)
Reading familiar words: % Accuracy	89.6 (6.9)	91.6 (8.0)
Reaction time in milliseconds for accurate responses.	1725.8 (570.7)	1058.7 (463.0)

Table 4). However, while rate of reading was significantly correlated with both comprehension and word accuracy in the non-phonics group, it was not significantly correlated with either in the phonics group.

In the non-phonics group the fastest readers are also those who best comprehend the story. This is understandable as the fewer breaks in reading the story, the less likely comprehension will suffer. In the phonics group, however, the children have been taught to spend time sounding out words they do not know. Hence all of these children take much time, which means that individual differences in their Rate of Reading do not have significant correlations with the other Neale measures. Reading comprehension is significantly correlated with aural vocabulary (BPVS score) in the non-phonics group, but not the phonics group.

Another explanation for the rate measures being uncorrelated may be the lower reliability of the Neale Rate measure and it's higher standard error of measurement compared with the Accuracy and Comprehension measures (Neale 1989).

Correlational analysis also showed that those who did well at the BAS word reading test are those who also did well at the Neale test and are good at nonword reading. Contextual cues were available to aid reading in the Neale tests but it was overwhelmingly those who were good at reading without context (i.e. high BAS word recognition score) that were also best at reading with context available (i.e. high Neale score).

Table 4. Correlations of Neale analysis of reading ability scores with reading and reading related measures by group

Measures	Neale accuracy	Neale comp.	Neale rate	BAS word test	Nonword reading	Phoneme segmen-tation	Word RT	Regular words	Exception words	BPVS scord
	Phonics group, n = 22									
Neale accuracy	–	0.68**	0.31	0.83**	0.62**	0.49*	–0.08	0.64**	0.66**	–0.01
Neale comp.		–	0.09	0.46*	0.35	0.07	–0.09	0.33	0.36	0.17
Neale rate			–	0.58**	0.43*	0.48*	–0.33	0.52*	0.47*	–0.00
	Non-phonics group, n = 26									
Neale accuracy	–	0.72**	0.75**	0.88**	0.63**	0.24	–0.26	0.75**	0.74**	0.25
Neale comp.		–	0.49*	0.67**	0.49*	0.16	–0.07	0.50*	0.46*	0.51**
Neale rate				0.69**	0.42*	0.12	–0.13	0.56**	0.49*	–0.05

Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$
Neale comp. = Neale comprehension measure, RT = reaction time (Negative correlations with Rate measures expected).

Yopp Singer Phoneme Segmentation Test, Yopp (1988)

Performance on this measure was higher for the phonics taught group than the non-phonics taught group [$t(46) = 2.75$, $p = 0.008$]. See Table 3. The phonics group performance on the Neale Word Accuracy and Reading Rate measures correlated significantly with performance on the Yopp-Singer task. This was not the case for the non-phonics group whose Yopp-Singer results showed no significant correlation with the Neale measures including Word Accuracy.

Regular and exception word reading

Analysis of variance on these results, which are displayed in Table 3, showed that there was a significant interaction of group (phonics and non-phonics) and stimuli (regular and exception) [$F(1, 44) = 15.8$, $p < 0.000$]. Post hoc analyses revealed that this was due to the non-phonics taught children reading exception words more accurately than the phonics taught children and in turn the phonics taught children correctly reading more of the regular words. However, the absolute difference in terms of actual words read was small and this result needs to be treated with some caution given the small number of stimuli used.

Reaction times for familiar words

As expected for familiar words very high accuracy scores are shown in Table 3 for the reaction time task. A *t*-test showed no significant difference between the two groups in terms of reading accuracy [$t(46) = 0.93$, $p = 0.360$]. In this task only reaction times to correct responses have been analysed so there are varying numbers of trials for each child. Reaction time data can often show a positively skewed distribution since it tends to magnify stimulus anticipation and lapses of attention. To minimise statistical artefacts, the geometric mean reaction time for each correct word was calculated for each child. This method was chosen for its ability to handle skewed data. The results in Table 3 show the arithmetic means of the geometric mean reaction times for each type of word. The arithmetic mean is used as it is expected that any skew in the data distribution will have been normalised by the use of the geometric means for the individual children.

A *t*-test on the reaction time data showed that the non-phonics group was significantly faster at naming the words than the phonics group [$t(46) = 4.29$, $p < 0.000$]. This again emphasises the idea that the phonics taught children were probably using explicit phonological recoding for the words, in spite of the fact that they were familiar classroom words. It was noted that just before the phonics taught children would overtly respond to a word they would

mouth their way through almost all the sounds of the word. This may explain the unusually long reaction times of the phonics taught children compared to the non-phonics taught children. The non-phonics taught children were fast readers compared to the first graders Lesgold, Resnick and Hammond (1985) tested whose mean reaction times were well above 1000ms.

Nonword reading

The phonics taught children were significantly more accurate than the non-phonics taught children at reading the 24 nonwords [$t(46) = 5.39$, $p < 0.000$]. See Table 3. This difference was large in spite of the words being orthographically similar to words with which the children were familiar. Nonword reading skill correlated significantly with performance on the Neale measures of word accuracy and reading rate for both groups of children.

Error analysis

The errors made by the children during the Neale reading and the first BAS Word Reading test were analysed. Previous studies (e.g. Elder 1971; Seymour & Elder 1986) found differences in error rates which were ascribed to instructional differences in reading approach. Table 5 gives the average percentage of error types as a proportion of total errors produced. The errors were classified into spoken response errors (which could be either real word responses or nonsense word responses to the test words and also included additions/insertion errors to the Neale texts) and refusal errors (i.e. when the child gave no response to the word being read after 5 seconds or said they did not know that word) for the BAS word reading test and the word responses in the Neale text reading. Omission errors were very rare in the Neale texts (only 5 in total across all the children) and were not analysed because of this and as no meaningful comparison could be made with the BAS error set.

There was no significant difference in amount of total raw errors produced by the phonics and non-phonics taught children for either the BAS Word Reading test [$t(46) = 1.43$, $p = 0.143$] or the Neale test [$t(46) = 1.25$, $p = 0.219$].

Analysis of variance was carried out on the refusal errors as percentages of the total errors, by group (phonics or non-phonics) and test (BAS test, Neale test). It was found that there was a significant main effect of group [$F(1, 45) = 29.9$, $p < 0.000$] but no significant interactions or other main effects ($F < 1$). The non-phonics taught children produced more refusal errors than the phonics taught children in both the BAS test and the Neale test. The non-phonics taught children were not so willing to try and attempt an unknown word and this was so even when they had context to help them as in the Neale

Table 5. Mean number of total errors and percentage of error types in the BAS word reading and Neale analysis of reading ability (standard deviations in parentheses)

	Phonics group		Non-phonics group	
	n = 22		n = 26	
	BAS	Neale	BAS	Neale
Mean Total number of errors	29.6 (5.9)	10.4 (5.5)	27.1 (5.5)	8.8 (5.4)
Total errors				
Percentage refusals	7.4 (16.7)	18.1 (22)	45.5 (35.4)	48.3 (26.1)
Percentage spoken responses	92.6 (16.7)	81.9 (22)	54.5 (35.4)	51.7 (26.1)
Totals	100	100	100	100
Spoken response errors:				
Nonsense words	48.4 (17.5)	34.3 (23.8)	29.6 (24.7)	16.8 (21.7)
Real words	51.6 (17.5)	65.7 (23.9)	70.4 (24.7)	83.2 (21.6)
Totals	100	100	100	100

text reading. It would appear that if the child is not willing to work out words, providing context will not alter that strategy. This interpretation was further supported by the positive correlation [$r(26) = 0.53$, $p < 0.05$] in the non-phonics group between refusals production in the BAS test and refusals in the Neale test. The phonics group showed no significant correlation between refusal production in the two tests [$r(21) = 0.08$, $p > 0.05$].

The spoken response errors were classified into real word and nonsense word responses (see Table 5 for the means). Analysis of variance on the percentages of nonsense word errors among spoken response errors revealed a significant main effect of group [$F(1, 45) = 12.1$, $p = 0.008$] and test [$F(1, 45) = 8.7$, $p = 0.01$] but not of group by test ($F < 1$). The phonics group produced a larger proportion of nonsense word errors among the spoken response errors than the non-phonics group in both tests. However, overall there were fewer nonsense word errors in the Neale test than in the BAS test for both phonics and non-phonics groups. This shows that context was having an effect on responses to unknown words for both groups, as nonsense words would not fit into the story context of the Neale test.

Contextually appropriate errors. In the Neale test, real word errors that fitted with either the immediate context of the sentence or story line of the each text were counted for each child. An error was counted as fitting the immediate

context of the sentence if it was judged that a meaningful sentence could be carried on by the addition of the error word to the preceding sentence text. For example, instead of "In the middle of the traffic lay two children" the error "In the middle of the traffic lay two mice." would be counted as contextually appropriate while the error "In the middle of the traffic lay two the." would not be contextually appropriate. An error that produced a word that fitted the preceding general context or theme of the story line was also counted as contextually appropriate. For example, instead of "She made a nest in my garden. Now I look after her little ones." the child read "She made a nest in my garden. Now I look after her little nest."

Our definition of contextually appropriate was very broad and some errors that fitted the immediate context of the sentence may not have fitted the story line and vice versa but were counted as contextually appropriate. However, since the children in the non-phonics regime were encouraged to use both the immediate sentence context and the story line context when working out unknown words we believe such a broad definition is justified.

It was hypothesised that the non-phonics taught children would be true to their teaching and use graphemic sources to identify words only after all other contextual information had been used. Therefore, it was thought that their contextually appropriate errors should contain less graphemically correct information than the phonics taught children whose first approach should be to sound out an appropriate word that fits the context. All the contextually appropriate errors were classified into those errors that had two or more letters correct and in the correct word position versus those that had less than two letters correct and in the correct word position.

The incidence of these errors as a percentage of the real word spoken response errors are reproduced in Table 6.

Analysis of variance showed no main effect of group (phonics vs non-phonics), [$F(1, 46) = 2.8$, $p = 0.102$] but did show a main effect of error type (with 2 or more letters matching versus less than 2 letters matching), [$F(1, 46) = 12.8$, $p = 0.01$]. There was also a significant interaction of group by error type [$F(1, 46) = 4.07$, $p = 0.049$]. Post hoc testing showed that the phonics taught children produced significantly more contextually appropriate errors that had two or more letters correct than the non-phonics taught children. There was no difference between the groups at producing contextually appropriate errors with less than 2 letters correct. This result shows that the phonics group produced a larger percentage of contextually appropriate errors with more intact graphemic information than the non-phonics taught children. The non-phonics group, despite their teaching that emphasises context, did not produce more contextually appropriate errors than the phonics group. The

Table 6. Contextually appropriate errors as a mean percentage of real word spoken response errors produced in the Neale analysis of reading ability (standard deviations in parentheses)

	Phonics group	Non-phonics group
Error type	n = 22	n = 26
Contextually appropriate errors	(a)	(b)
– with 2 or more letters matching in correct position in word	57.7 (27.1)	36.9 (32.1)**
– with less than 2 letters matching in correct position in word	20.8 (24.2)	26.6 (25.5)N.S.

** = Significant difference between column a and column b.
N.S. = No significant difference between column a and column b.

phonics group are using word and context information to identify words in text.

Discussion

This study has shown that when matched on word recognition phonics taught beginning readers were better at reading comprehension than children not receiving phonics. The type of processing used for word recognition differed between the two groups of beginning readers. The phonics group was superior to the non-phonics group in both nonword reading and phonemic awareness. They were also superior in reading words with regular letter-sound correspondences but inferior to the non-phonics group in reading words with exceptional correspondences. All these findings point towards the phonics taught children having greater proficiency in explicit phonological recoding.

However, the phonics taught children were much slower in reading text (Neale text) than the non-phonics taught children; also reading reaction time to familiar words was much longer for the phonics taught children. These results are similar to those reported for beginners with and without phonics by Lesgold, Resrick and Hammond (1985). Lesgold et al. (1985) found that their non-phonics sample was reading 40–45 words per minute during the equivalent period at school. Their phonics sample had begun to rapidly increase in speed at the end of Grade 1 and so were also reading at about 40–45 words per minute. Therefore, our phonics sample was about $\frac{1}{4}$ of a Grade behind in the number of words read per minute while our non-phonics sample was very similar to the Lesgold et al. samples in number of words read per minute.

Our two samples were faster at responding to individual words than the Lesgold et al. (1985) samples at the equivalent points in the school year. The phonics taught children in the Lesgold et al. study averaged about 2.5 seconds to respond to individual words while the non-phonics taught children averaged about 2.2 seconds. This difference may be explained by the low accuracy rates of the Lesgold et al. samples. The Lesgold et al. phonics taught sample averaged only 52% correct and the non-phonics taught sample only 65% correct. Therefore the word set presented by Lesgold et al. was more difficult than the word set we presented to our young readers and this may have slowed down the Lesgold et al. readers. However, the general pattern of results Lesgold et al. found is still very similar to our results with non-phonics taught children faster than phonics taught children.

The error analysis for text reading showed that the phonics taught children were spending time making spoken attempts at over 80% of unknown words, whereas the non-phonics taught children spent such time on little more than 50% of unknown words. This also applied to reading of isolated words (BAS word reading). So, unexpectedly, the non-phonics group, who received teaching with much emphasis on the children using context for recognition of words, were not more advantaged than the phonics group in attempting more responses to unknown words in text. In both groups, individual differences in rate of text reading were not significantly related to the children's reaction times in reading familiar words. The differences in text reading were due to time spent on attempts at explicit phonological recoding of unknown words. This is indicated by the fact that a third of the spoken attempts at unknown words of the phonics taught children were nonsense words but only a sixth among the non-phonics taught children.

In spite of the phonics group being much slower in their rate of reading text, they were better than the non-phonics group in answering questions on the text at the conclusion of their reading. This may seem surprising, as it has been claimed that speeding young children's pacing of reading will improve reading comprehension (Breznitz & Share 1992), and that when reading rate is too slow, comprehension processes and the facilitative effect of context on word recognition both suffer (Smith 1978). However, there is evidence against this second claim, at best with respect to the facilitative effect of appropriate context on word recognition (Thompson 1981). In the results of the present study, we have evidence against the claim that a slow rate of reading disrupts the child's use of context for word recognition, at least among attempts at unknown words. In fact, the phonics group which was reading at a slow rate were using context for this purpose proportionately more than the faster non-phonics group. However, it is wise to note that our sample of children were very young and at the early stages of reading. It is

presumed that they will become much more fluent with time and experience of reading. Rate of reading will become more important once fluent word recognition skills have been established.

Why is the comprehension of the slow reading phonics taught children better? When they came across an unfamiliar word they tried to work it out by sounding it out, in order to fit their response appropriately to the story context. This procedure is indicated by the greater proportion of contextually appropriate errors produced by the phonics taught children which are graphemically similar to the target word, and the evidence about their greater use of explicit phonological recoding. This procedure may encourage them to think about the story as they sound out words and check that their responses match the context. The procedure would lead to greater rehearsal of content of the story as reading proceeds. Hence, although the phonics taught children are slower in their reading they comprehend and remember more about the story.

The non-phonics taught children produce more refusals when trying to work out unfamiliar words. This gives them fewer opportunities to rehearse the content of the story as their reading of the text proceeds. Those spoken errors the non-phonics taught children did make were less likely to fit the context than those produced by the phonics taught children. This may be surprising given the emphasis that the teaching they received had on the use of context for word recognition. But it is not surprising if they do not retain the context of the story so effectively as they are reading the text. It is also relevant to observe that among the phonics group, individual differences in reading comprehension were not significantly correlated with aural vocabulary, but they were in the non-phonics group. The non-phonics group appeared to be relying more on general verbal ability to achieve their more limited comprehension of the texts.

Skill at using phonological recoding and context to read unfamiliar words has been described by Share (1995) as a self-teaching mechanism for acquiring reading vocabulary. Share sees this self-teaching mechanism as the way that young children develop their word reading skills beyond the initial, inflexible, and limited sight vocabulary of the very inexperienced reader. Recent results of Tunmer and Chapman (1998, 1999) support this view. Our results would also support such a view in which explicit phonological recoding and context were used in concert in the phonics taught sample. Among them the mechanism provides a small but significant advantage for reading comprehension early in learning to read. This advantage is above and beyond that of any advantage due to better word recognition skill.

It should be noted that the non-phonics taught children were not "logographic readers" by any current definition (see Ehri 1992; Frith 1985;

Seymour & Elder 1986; Share 1995; Share & Stanovich 1995). They were able to read some nonwords successfully and they produced some nonsense word errors in word and text reading. Their word reading correlated positively with nonword reading. However, there was no significant correlations between any of the Neale reading measures and the phonological awareness measure, unlike the phonics taught group. This pattern of correlations could be explained by the fact that the phonological awareness task closely fits the type of explicit phonological recoding taught to, and used by, the phonics group but not by the non-phonics group, who apparently may only have a more implicit form of phonological recoding available (Thompson & Johnston, in press).

Our two groups do not differ in real word reading accuracy but differ in a test of phonological segmentation The difference, however, between the groups in absolute terms is small, being only of the order of 2.5 items and should be interpreted with caution. Many other studies have not found such a difference. Thompson and Johnston (In press), for example, found no difference in phonemic awareness between matched samples of phonics and non-phonics taught children in children of the same age but using a more stringent phoneme deletion test. Other studies have found similar results (Foorman, Novy, Francis & Liberman 1991; Hatcher, Hulme & Ellis 1994; Iversen & Tunmer 1993; Tunmer & Nesdale 1985). What may be more important is not the absolute level of phonemic awareness skill shown by these tests but how phonemic awareness is used by the children when learning to read. The phonics taught children's reading scores correlate significantly with phonemic awareness while the non-phonics taught children's scores do not.

It is difficult and probably premature to draw conclusions about instructional recommendations from the small sample of readers described here (not including those at the upper or lower ends of the normal distribution). It is clear, however, that heavy phonics instruction in the early years of schooling does not prevent the development of age appropriate comprehension skills in reading. In fact, it appears that early imperfect attempts to phonologically recode text encourage the use of contextual cues in order to help select an appropriate sounding word that will make sense in the story being read.

Conclusions. A phonics taught sample of beginning readers were compared with a sample without phonics teaching, but at the same level of word recognition accuracy. The phonics taught children spent more time in attempts at identifying unknown words, using both phonological recoding and contextual information. This additional time in using contextual information apparently resulted in more rehearsal of the content of the text, and hence a small but

significant advantage to the phonics taught children in their comprehension of the texts, which they read much more slowly than the children without explicit phonics teaching. The texts used here included those which were very difficult for the beginning reader, and there was only a very small sample of the texts. It should not be assumed that these results would also apply to the effects of phonics instruction on beginning reading of easy texts. This is a question for further research.

Acknowledgments

We would like to thank the children and staff of the schools in Tayside, Scotland and Wellington, New Zealand for their co-operation throughout these studies. We also thank the three reviewers of this paper who provided stimulating and important commentary on our text.

Appendix A

Exception and Regular word stimuli (from Holligan & Johnston 1991)

Exception	*Regular*
Sword	luck
Pint	spear
Broad	mile
Lose	beer
Wool	stuck
Come	door
Foot	stick
Shall	best
Both	green
Great	fine

Appendix B

Familiar classroom word stimuli for the reading reaction time task			
go	you	came	they
he	one	like	play
me	too	home	this
no	are	make	what
we	the	here	stop
am	but	look	down
at	can	said	with
in	got	good	went
on	for	read	help
up	not	rain	bird

Appendix C

Nonword stimuli			
ga	ome	dith	rike
ap	bup	plam	loof
na	tob	hent	hobe
ig	caz	thit	sood
er	arg	lant	pake
ob	lor	stom	tain

References

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

Barr, R. (1975). The effect of instruction on pupil reading strategies, *Reading Research Quarterly* 10: 555–582.

Beck, I.L. & Block, K.K. (1979). An analysis of two beginning reading programs: Some facts and some opinions. In: L.B. Resnick & P.A. Weaver (eds.), *Theory and practice of rarly teading*, Vol. 1 (pp. 279–318), Hillsdale, NJ: Lawrence Erlbaum Publishers.

Bond, G.K. & Dykstra, R. (1967). The co-operative research program in first grade reading instruction, *Reading Research Quarterly* 2: 5–142.

Breznitz, Z. & Share, D.L. (1992). Effects of accelerated reading rate on memory for text, *Journal of Educational Psychology* 84(2): 193–199.

Carroll, J.B., Davies, P. & Richman, B. (1971). *The word frequency book*. New York: American Heritage.

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

Clay, M.M. (1979). *The early detection of reading difficulties*. London: Heinemann.

Curtis, M.E. (1980). Development of components of reading skill, *Journal of Educational Psychology* 72: 656–669.

Dunn, L.M. & Dunn, L.M. (1982). *British picture vocabulary scale*. Windsor: NFER-Nelson.

Ehri, L.C. (1992). Reconceptualizing the development of sight word reading and its relationship to decoding. In: P.B. Gough, L.C. Ehri & R. Treiman (eds.), *Reading acquisition* (pp. 107–143), New Jersey: Erlbaum.

Elder, R.D. (1971). Oral reading achievement of Scottish and American children. *Elementary School Journal* 71: 216–230.

Elliott, C.D., Murray, D.J. & Pearson, L.S. (1977). *The British ability scales*. Windsor: NFER-Nelson.

Foorman, B.R., Novy, D.M., Francis, D.J. & Liberman, D. (1991). How letter sound instruction mediates progress in 1st grade reading and spelling, *Journal of Educational Psychology* 83(4): 456–469.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In: Patterson, K.E., Marshall, J.C. & Coltheart, M. (eds.), *Surface dyslexia* (pp. 301–330), London: Erlbaum.

GINN (1988a). *Reading 360 levels 1 and 2 teachers resource book*. Aylesbury, UK: GINN & Company Ltd.

GINN (1988b). *Reading 360 levels 3 and 4 teachers resource book*. Aylesbury, UK: GINN & Company Ltd.

Glynn, T., Crooks, T., Bethune, N., Ballard, K. & Smith, J. (1989). *Reading recovery in context*. Wellington, New Zealand: Department of Education.

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(1): 41–57.

Holligan, C. & Johnston, R.S. (1991). Spelling errors and phonemic segmentation a bility: The nature of the relationship, *Journal of Research in Reading* 14: 21–32.

Iversen, S. & Tunmer, W.E. (1993). Phonological processing skills and the reading recovery programme, *Journal of Educational Psychology* 85(1): 112–126.

Johnston, R.S. & Thompson, G.B. (1989). Is dependence on phonological information in childrens reading a product of instructional technique? *Journal of Experimental Child Psychology* 48: 131–145.

Lesgold, A., Resnick, L.B. & Hammond, K. (1985). Learning to read: A longitudinal study of word skill development in two curricula. In: G.E. MacKinnon & T.G. Waller (eds.), *Reading research: Advances in theory and practice*. Vol. 4 (pp. 107–138), New York: Academic Press.

Neale, M.D. (1989). *Neale analysis of reading ability*, Revised British Edition. Windsor, Berkshire, UK; NFER-Nelson.

New Zealand Department of Education (1985). *Reading in junior classes (With guidelines to the revised ready to read series)*. Wellington, New Zealand: Government Printer.

Perfetti, C.A. (1985). *Reading ability*. New York: Oxford University Press.

Seymour, P.H.K. & Elder, L. (1986). Beginning reading without phonology, *Cognitive Neuropsychology* 1: 43–82.

Seymour, P.H.K. & Evans, H.M. (1992). Beginning reading without semantics: A cognitive study of hyperlexia, *Cognitive Neuropsychology* 9: 89–122.

Share, D.L. (1995). Phonological recoding and self teaching: sine qua non of reading acquisition, *Cognition* 55: 151–218.

Share, D.L. & Stanovich, K.E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition, *Issues in Education* 1: 1–58.

Smith F. (1978). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Holt, Rhinehart and Winston.

Sowden, P.T. & Stevenson, J. (1994). Beginning reading strategies in children experiencing contrasting teaching methods, *Reading and Writing* 6: 109–123.

Stanovich, K.E. (1980). Towards an interactive-compensatory model of individual differences in the development of reading fluency, *Reading Research Quarterly* 16: 32–71.

Stanovich, K.E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity and modularity, *Developmental Review* 10: 72–100.

Stuart, M. & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition* 30: 139–181.

Thompson, G.B. (1981). Semantic context and graphic processing in the acquisition of reading, *British Journal of Educational Psychology* 51: 291–300.

Thompson, G.B. (1982). Initial vocabularies for reading. *Set: Research information for teachers*. Australian Council for Educational Research. No. 2, Item 3.

Thompson, G.B. (1993). Reading instruction for the initial years in New Zealand Schools. In: G.B. Thompson, W.E. Tunmer & T. Nicholson (eds.), *Reading acquisition processes* (pp. 148–154), Clevedon, England: Multilingual Matters Ltd.

Thompson, G.B. & Johnston, R.S. (1993). The effects of type of instruction on processes of reading acquisition. In: G.B. Thompson, W.E. Tunmer & T. Nicholson (eds.), *Reading acquisition processes* (pp. 74–90), Clevedon, England: Multilingual Matters Ltd.

Thompson, G.B. & Johnston, R.S. (in press) Are nonword and other phonological deficits indicative of a failed reading process? *Reading and Writing* (in press).

Tunmer, W.E. & Chapman, J.W. (1998). Language prediction skill, phonological recoding ability and beginning reading. In: C. Hulme & R.M. Joshi (eds.), *Reading and spelling: Development and disorder* (pp. 33–67), Hove, UK: Erlbaum.

Tunmer, W.E. & Chapman, J.W. (1999). Teaching strategies for word recognition. In: G.B. Thompson & T. Nicholson (eds.), *Learning to read: Beyond phonics and whole language* (pp. 74–102), New York: Teachers College Press, Columbia University.

Tunmer, W.E. & Nesdale, A. R. (1985). Phonemic segmentation skill and beginning reading, *Journal of Educational Psychology* 77(4): 417–427.

Vellutino, F.R. (1991). Introduction to three studies on reading acquisition, *Journal of Educational Psychology* 83: 437–443.

Venezky, R.L. (1973). The letter sound generalisations of first, second and third grade Finnish children, *Journal of Educational Psychology* 64: 288–292.

Yopp, H.K. (1988). The validity and reliability of phonemic awareness tests, *Reading Research Quarterly* 23: 159–177.

Wimmer, H. & Hummer, P. (1990). How German speaking first graders read and spell: doubts on the importance of the logographic stage, *Applied Psycholinguistics* 11: 349–368.

Address for correspondence: Vincent Connelly, Psychology, Erlang House, South Bank University, 103 Borough Road, London SE1 0AA, United Kingdom
Phone: (+44) 171 815 8488; Fax: (+44) 171 815 8099; E-mail: connelv@sub.ac.uk