Australasian Journal of Special Education
Vol. 32, No. 1, April 2008, pp. 23–54

Corrective Reading: An Evidence-Based Remedial Reading Intervention

Kerry Hempenstall*
RMIT University, Bundoora, Australia

This article first examines recent theoretical and empirical research on reading development and instruction in English-speaking countries. Then, a study is described that examines the effects of a synthetic phonics-emphasis Direct Instruction remedial reading program on the phonological processes of students, with teacher-identified serious reading problems, attending several Melbourne suburban schools. The 206 students (150 males and 56 females, mean age 9.7 years) were pre-tested on a battery of phonological tests, and assigned to the treatment condition or to a wait-list comparison group. The 134 students in the intervention group received the 65 lessons (in groups of up to 10) of the *Corrective Reading: Decoding* program from reading teachers at their schools. When compared with a similar cohort of 72 wait-list students from the same schools, the students made statistically significant and educationally large gains in the phonologically-related processes of word attack, phonemic awareness, and spelling, and statistically significant and moderately large gains in phonological recoding in lexical access, and phonological recoding in working memory. The study contributes to an understanding of the relationship between phonological processes and reading, and to an approach to efficiently assisting students whose underdeveloped decoding places their educational progress at risk.

The Direct Instruction model has a relatively long history in reading education, the first program having been published in 1969. However, there has been surprisingly little serious attention paid to it from both the educational bureaucracy and the educational research community, despite its strong body of supportive empirical evidence. Reports of Operation Follow Through (Engelmann, Becker, Carnine, & Gersten, 1988; Grossen, 1996), and the studies reported in meta-analyses by White (1988) and by Adams and Engelmann (1996) have not been accorded the attention that might have been expected. However, this anomaly has been part of a long lamented and broader malaise—the failure of research-based knowledge to have an impact upon educational decision making (Carnine, 1995; Hempenstall, 1996, 2006; Stanovich, 1994, Stone, 1996).

In recent years, there has been a marked change in this education myopia in the USA and Great Britain, though, as yet, not in Australia. It is a change that is also

*Division of Psychology, School of Health Sciences, RMIT University, Plenty Rd., Bundoora, Victoria, Australia. 3083. Email: kerry.hempenstall@rmit.edu.au

ISSN 1030-0112 (print)/ISSN 1833-6914 (online)/08/010023-32
© 2008 Australian Association of Special Education Inc
DOI: 10.1080/10300110701839949

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

evident in fields other than reading instruction, for example, the rise of *evidence-based medicine* in patient care (Sackett, Rosenberg, Muir Gray, Haynes, & Richardson, 1996), and *empirically validated treatment* in psychotherapy (American Psychological Association, 1993). These changes have been wrought despite significant resistance from entrenched traditionalists in their respective professions.

Evoking a sense of optimism is the gradual pressure for change spreading across those nations using written alphabetic languages. The similarity of recommendations among numerous national and state reports in the USA, for example, those of the National Institute of Child Health and Human Development (Lyon, 1998), the National Reading Panel (National Reading Panel, 2000), the American Institutes for Research (1999), the National Research Council (Snow, Burns, & Griffin, 1998), and the Texas Reading Initiative (1996), has demonstrated the considerable consensus existing about the crucial elements of reading development and instruction. The importance to successful instruction of phonemic awareness and the alphabetic principle has been strongly asserted. The extra traction gained by systematic synthetic phonics instruction over more ad hoc, loosely specified phonics approaches is clearly noted in these reports. The impact on students of such careful explication of the code can be described as inoculative against reading failure. In this sense, it parallels the use of a vaccine to evince immunity to a specific disease—a public health measure considered worthwhile for all, even though only some of the population may be at risk.

In Britain, the National Literacy Strategy (Department of Education and Employment, 1998) was released to all primary schools, requiring them to abandon the Whole Language approach to reading. Components of the former system, such as teaching students to rely on context clues to aid word reading, were discredited in the Strategy, and explicit phonics instruction was mandated from the earliest stages of reading instruction. 'There must be systematic, regular, and frequent teaching of phonological awareness, phonics and spelling' (Department for Education and Employment, 1998, p. 11). Unfortunately, the strong resistance to such explicit teaching led to substantially less instructional change than was anticipated, and correspondingly less improvement in national literacy figures. In 2006, the Primary Framework for Literacy and Mathematics (Primary National Strategy, 2006) was released, updating its 1998 predecessor, and mandating practice even more firmly onto an evidence base. In particular, it withdrew its imprimatur from the 3-cueing system (Hempenstall, 2003) approach to reading, and embraced the Simple View (Hoover & Gough, 1990) of reading that highlights the importance of decoding as the pre-eminent strategy for *saying what's on the page.* Under the 3-cueing system, making meaning by any method (pictures, syntactic, and semantic cues) took precedence over decoding as the prime strategy. The new 2006 Strategy mandates a synthetic phonics approach, in which letter–sound correspondences are taught in a clearly defined sequence, and the skills of blending and segmenting phonemes are assigned high priority. This approach contrasts with the less effective analytic phonics, in which the phonemes associated with particular graphemes are not pronounced in isolation (i.e., outside of whole words). In

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

the analytic phonics approach, students are asked to analyse the common phoneme in a set of words in which each word contains the phoneme being introduced (Hempenstall, 2001). The lesser overall effectiveness of analytic phonics instruction may be due to a lack of sufficient systematic practice and feedback usually required by the less able reading student (Adams, 1990).

In Australia, the National Enquiry into the Teaching of Literacy (Department of Education, Science, and Training, 2005) recommendations exhorted the education field to turn towards science for its inspiration. For example, the committee argued strongly for empirical evidence to be used to improve the manner in which reading is taught in Australia.

> In sum, the incontrovertible finding from the extensive body of local and international evidence-based literacy research is that for children during the early years of schooling (and subsequently if needed), to be able to link their knowledge of spoken language to their knowledge of written language, they must first master the alphabetic code—the system of grapheme–phoneme correspondences that link written words to their pronunciations. Because these are both foundational and essential skills for the development of competence in reading, writing and spelling, they must be taught explicitly, systematically, early and well. (p. 37)

Perhaps extra impetus for similar reform in Australia will arise from the *Parents' Attitudes to Schooling* report from the Department of Education, Science and Training (2007). Among the findings was that only 37.5 per cent of the surveyed parents believed that students leave school with adequate literacy skills. Generally, the impact of state and national testing has led to greater transparency concerning how our students fare in their literacy development. Media attention on these findings and on the occasional litigation have focused community attention, and (thereafter) renewed government attention to the issue of reform.

From a theoretical perspective, each of the National Reading Panel (2000) recommended foci for reading instruction (phonemic awareness, phonics, fluency, vocabulary, comprehension) is clearly set out and taught in Direct Instruction literacy programs. These same instructional features were endorsed in the report of the National Enquiry into the Teaching of Literacy. An examination of the program teaching sequences in, for example, the Reading Mastery (Engelmann & Bruner, 1988) and Corrective Reading (Engelmann, Hanner, & Johnson, 1999) texts attests to their comprehensive nature.

However, these necessary elements are only the ingredients for success. Having all the right culinary ingredients doesn't guarantee a perfect soufflé. There are other issues, such as what proportion of each ingredient is optimal, when should they be added, how much stirring and heating is necessary? Errors on any of these requirements lead to sub-optimal outcomes.

So it is with literacy programs. 'Yet there is a big difference between a program *based* on such elements and a program that has itself been compared with matched or randomly assigned control groups' (Slavin, 2003, p. 15). Just because a program has

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

most, or all, of the elements doesn't guarantee that it will be effective necessarily. Engelmann (2004) points to the logical error of inferring a whole based upon the presence of some or all of its elements. The logic error is seen in the following: *If a dog is a Dalmatian, it has spots. Therefore, if a dog has spots, it is a Dalmatian* (Engelmann, 2004, p. 34). In this simile, the Dalmatian represents programs known to be effective with students. It is possible to analyse the content of these programs, and then assume incorrectly that the mere presence of those characteristics is sufficient to ensure effectiveness. This ignores the orchestration of detail that also helps determine effectiveness. Engelmann is thus critical of merely 'research-based' programs, that is, programs constructed only to ensure each respected component is somewhere represented in the mix.

Reading First is a massive program in the USA designed to improve literacy outcomes for disadvantaged students in the first four years of schooling. Early reports (Office of Management and Budget, 2007) indicate that it is having a positive impact nationally; however, one criticism of it is that the criterion for acceptability of the programs used was diluted. Reid Lyon, the primary architect of Reading First, was critical of the modification of his plan that funding should be provided only for programs with proven effectiveness—to the easier-to-meet criterion that programs had only to be *based on* scientifically based reading research (Shaughnessy, 2007). A possible reason for this Department of Education decision relates to the lack of well-designed studies of reading instruction. According to Slavin (2007), there are only two beginning programs generally acknowledged to have strong empirical evidence of effectiveness: Success for All and Direct Instruction. It was considered politically unacceptable to allow only two programs to dominate beginning reading to the nation's disadvantaged children. This decision has other ramifications. It has led to some programs offering only the appearance of being evidence-based, thereby diminishing the potential of the national scheme overall.

In most published reading schemes, program designers assume that teachers know how to structure a lesson effectively when they are provided with some worthwhile content. This assumption is far from universally justified. The content may be research-based, but its presentation may be competent, slipshod, or cursory. Corrective feedback may or may not occur systematically. Mastery by students may or may not be expected. Practice opportunities may or may not be adequate for the population. Regular data-based monitoring may or may not occur. Teacher creativity may abound. This loose coupling between content and delivery would horrify an empirically-trained psychologist, as it would a surgeon trained to follow protocols. It also highlights why the crucial element in evaluation is not simply that a program is consistent with scientific findings, but also that it has been demonstrably successful with the target population.

So for a true measure, we must look beyond theoretical acceptability, and examine empirical studies to show that a particular combination of theoretically important elements is indeed effective. The questions become: Has a particular program demonstrated independently replicated effectiveness? For what populations?

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

The development of criteria for acceptable research evidence is a common element in the reweighting of empirical research in the professional fields mentioned earlier. In the case of reading, it should make easier the task of convincing the educational community how valuable could be the findings of rigorous research in informing practice. Having established these criteria, it becomes easier to determine which of the plethora of reading programs available do have adequate research support at any given time. Unfortunately, the standard of educational research generally has not been high enough to enable confidence in its findings. Partly, this is due to a preponderance of short-term, inadequately designed studies. When Slavin (2004) examined the *American Educational Research Journal* over the period 2000–2003, only 3 out of 112 articles reported experimental/control comparisons in randomised studies with reasonably extended treatments.

The examination of existing evidence employing criteria (of various levels of stringency) by a range of groups has supported Direct Instruction as a valuable approach to reading instruction for both regular and struggling readers. For example, the American Federation of Teachers series of documents *Building on the best: Learning from what works* (1997) nominates Direct Instruction programs among each of its recommendations across different facets of education: *Seven Promising Reading and English Language Arts Programs, Three Promising High School Remedial Reading Programs, Five Promising Remedial Reading Intervention Programs* and *Six Promising Schoolwide Reform Programs.*

A report from the American Institutes for Research (1999), *An educators' guide to school-wide reform*, found that only three programs, Direct Instruction among them, had adequate evidence of effectiveness in reading instruction. In a follow-up evaluation (American Institutes for Research, 2006), 800 studies of student achievement were reviewed involving 22 programs directed at US high-poverty, low-performing schools. The two programs rated most highly were those that offered a high level of manualisation of both curriculum and non-curriculum features. The level of detail and the field testing and rewriting that occur before these programs are published does not preclude excursions from fidelity, but on average it does attenuate them. The two programs were Engelmann's Direct Instruction and Robert Slavin's Success for All.

Other similarly supportive reviews of Direct Instruction include: *Reading programs that work: A review of programs for pre-kindergarten to 4th grade* (Schacter, 1999), *Current practice alerts* (Council for Exceptional Children, 1999), *Bringing evidence driven progress to education* (Coalition for Evidence-Based Policy, 2002), *Center for Education Reform: Best bets* (McCluskey, 2003), *Comprehensive school reform and student achievement: A meta-analysis* (Borman, 2007; Borman, Hewes, Overman, & Brown, 2002), *Review of comprehensive programs* (Curriculum Review Panel, 2004), and *CSRQ Center Report on Elementary School CSR Models* (American Institutes for Research, 2005).

These reports have been influential in drawing attention to the large corpus of supportive research developed over the years indicative of the effectiveness of the

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Direct Instruction model across a wide range of educational settings. The model is now being implemented with varying degrees of fidelity in increasing numbers of school settings.

Considering the two aspects of reading research described above—the theoretical and the empirical—it is evident that the Direct Instruction model has strengths in each area to support its use. In line with current research findings, the programs focus on critical areas such as phonemic awareness (the ability to decompose the spoken word into its constituent sounds) and letter–sound relationships. The themes critical for struggling students are paid careful attention in the program design. These are adequate lesson frequency (daily), and sufficient daily and spaced practice to reduce the risk of forgetting, immediate correction of errors to guide the student towards mastery, and continuous assessment of progress to validate the effectiveness of the teaching. Refreshingly, the assessment emphasises the teaching process rather than the child as the major issue. Failure to learn is viewed as failure to teach effectively, and specific corrective teaching procedures are developed to redress the problems should lack of progress be observed. The emphasis on teaching quality rather than learner quality makes redundant any explanations of failure based on intelligence, race, readiness, first language, or home background. It is an empowering approach because it acknowledges and reinforces the status and power of teachers to make a real difference to students.

Interestingly, in Australia there has been a rise in the adoption of Direct Instruction programs without any state or federal government support until recently. About 350 schools in Victoria have implemented one or more Direct Instruction programs (McGraw Hill, personal communication, June 2007) across basic skill areas, such as language, reading decoding and reading comprehension, spelling, writing, and maths.

In state education department documents, the former wholesale acceptance of the Whole Language model has been sharply declining, except for the maintenance of a near-relation, Reading Recovery, as the first line of remediation. It is an expensive intervention, given that it is required by 40–50% of first grade students in Victoria (Office of the Victorian Auditor General, 2003). The recent review by Reynolds and Wheldall (2007) highlights the limitations of that approach in attempting to achieve universal literacy.

In recent times, there has been some interest in Direct Instruction from the federal government—*Working out what works* (Hoad, Munro, Pearn, Rowe, & Rowe, 2005), and in the literature review presented to the National Inquiry into the Teaching of Literacy: *A review of the empirical evidence identifying effective interventions and teaching practices for students with learning difficulties in Years 4, 5 and 6* (Purdie & Ellis, 2005). At the state level in Victoria, the *Successful Interventions Literacy Research Project* (Department of Education, Employment, and Training, 2001) reported favourably upon one such program—the Corrective Reading program.

Many of the schools employing Direct Instruction programs have opted for the remedial decoding program known as *Corrective Reading: Decoding* (Engelmann,

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Hanner, & Johnson, 1999) with mid-primary and older remedial readers, as it is around that time that the developmental lag explanation for a lack of a student's progress begins to ring hollow.

Because the *Corrective Reading* program has been available in various editions since 1978, there have been a significant number of evaluations completed, though not many have been sufficiently well designed to meet stringent publication criteria. Grossen (1998) reviewed some of the available research on the program, both controlled comparisons and school evaluations. She described eight studies that evaluated only Corrective Reading Decoding, one that evaluated only the sister program, Corrective Reading: Comprehension, and five that used both programs. The populations included general education students, limited English-speaking students, and special education students with various identified disabilities. Grossen reported that students in the Corrective Reading interventions progressed faster than students in the comparison groups in all but one of the studies. Similar results were reported for secondary students by Harris, Marchand-Martella, and Martella (2000), and also by Grossen (2004) in a larger-scale implementation. In a study of 45 incarcerated adolescents, Malmgren and Leone (2000) noted significant gains in fluency and accuracy of reading when 30 of 65 lessons of Corrective Reading: Decoding and Corrective Reading: Comprehension were provided. They also highlighted the problem of late intervention in observing that even these improved reading scores remained within the first percentile.

Sometimes the Direct Instruction programs have been modified for specific purposes. In a randomised design, Trezek and Malmgren (2005) successfully employed Decoding Level A, along with a means of making the articulatory gestures visual, with hearing impaired students. When compared to a comparison group, strong and significant differences were noted at post-test on identifying sounds in isolation, and on nonsense word reading.

Lovett et al. (1994) used a 35 lesson training program developed from *Reading Mastery: Fast Cycle 1 & 2* (Engelmann, & Bruner, 1984), and *Corrective Reading* to teach word identification to dyslexic students for one hour four times per week. They compared results to a control group taught a study skills program, and achieved highly significant post-test gains for the experimental group—effect sizes (*d*) of 0.76, 1.11, and 0.90 on the three training measures. The transfer to real words was impressive, and 'was based on the successful training of what is considered the core deficit of developmental dyslexia: phonological processing and nonword reading skill' (p. 818). Further, they argue, 'this training success rests on embedding letter–sound training in an intensive phonological training program' (p. 819). In a further study, offering 70 hours of Direct Instruction-based phonological instruction, Lovett et al. (2000) noted similarly large treatment effects, evident even in comprehension tasks.

In summary, Corrective Reading program's instructional content and design is considered to meet the criteria for acceptance as a scientifically based reading program (Oregon Reading First, 2004). However, there remains a need for better-quality studies to add to the research base (Smith, 2004).

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Another issue relevant to this study is the question: To improve decoding, is instructional time for struggling readers most effectively devoted to a dedicated phonemic awareness program along with a synthetic phonics program? Alternatively, is this precious time better spent solely in synthetic phonics activities? Johnston and Watson (2004) assert that phonological awareness training may be important alongside analytic phonics, but unnecessary when synthetic phonics is employed. In the Ehri et al. (2001) meta-analysis it is clear that the impact of phonemic awareness activities on subsequent reading is markedly enhanced when letters are part of the program. Hatcher, Hulme, and Ellis (1994) concluded that children in dual-input programs demonstrate more improvement in reading and spelling than those exposed to a solely oral phonemic awareness program. Hammill (2004) concurs, suggesting that a focus on print activities is more beneficial than on any other non-print activities. The findings of Bentin and Leshem (1993) suggest that for most children effective synthetic phonics programs are sufficient to evoke phonemic awareness alongside reading progress in beginning readers.

An interesting question is whether the programs can do so for older students with significant reading problems. Further, what about the other phonological processes? Does weakness in one or other of them need to be addressed specifically, or are they also amenable to improvement if reading itself can be developed?

This research was designed to assess the effect of participating in the *Corrective Reading* program on phonological processes (i.e., phonemic awareness, phonological recoding in lexical access, and phonological recoding in working memory), word attack, and spelling.

Method

I am often contacted by schools for advice on problems they may experience in effectively promoting student literacy. In the primary school setting, this most frequently involves students in middle and upper primary grades who appear to have experienced what Jean Chall referred to as the fourth grade slump (Rosenshine, 2002). For this study, I agreed to assist schools to establish a systematic synthetic phonics approach to address this predictable and regular problem. After receiving research-based information about the role of decoding deficits in the struggles students may have with reading success, a number of schools elected to take up this proposal. I agreed to provide ongoing program advice as needed, and to organise for pre-test and post-test assessment of all the students over the 65-lesson intervention.

The participants were 206 (150 male and 56 female) middle and upper primary school students attending five State and four Catholic schools in suburban Melbourne. A larger cohort of students had been referred for assistance by their class teachers because of perceived slow reading progress. The ratio of boys to girls in the larger cohort identified by teachers was similar to the final sex ratio in the study. Each of these students was individually assessed with the *Corrective Reading: Decoding* program *Placement Test* to ensure the presence of the program entry skills and the

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

absence of the program outcome skills. The final 206 participants comprised both the experimental (134 students) and the wait-list control group (72 students). The two groups constituted those referred students falling within the skill band suitable for inclusion in the *Corrective Reading: Decoding* program.

The experimental group consisted of students from those schools that had sufficient numbers of suitable students to form a group (about 10 students), and adequate school arrangements (e.g., staff) to enable the group(s) to proceed. Some schools had identified more students than they could manage at the one time. They elected for these students to delay participation in the intervention until the first group concluded the program. These latter students comprised the wait-list control list. They were receiving the regular grade-level English or reading program while the experimental group students were withdrawn for about 50 minutes—ideally five times per week. Hence the contrast was between two distinct interventions—the schools' English/reading programs and the Corrective Reading program.

There was no systematic allocation process that might be expected to produce different experimental/control group characteristics, and thus compromise the conclusions through selection bias. For example, there was no decision to intervene with the most delayed students first. When a school had both an intervention and a wait-list group, selection into intervention was based upon administrative criteria, such as being from the same grade. Additionally, the comparison groups were drawn from the same set of schools participating in the reading program; thus, there was less chance that socio-economic or other differences might confound the interpretation of results. The control group cluster may be described as non-equivalent (Cook & Campbell, 1979) as the students were not randomly assigned to their respective groups, but were from convenience samples.

More than 50% of the students were from areas considered disadvantaged. The low mean index (995) corresponds to the 25th percentile, indicating that the study areas have a high proportion of low-income families (Castles, 1994).

The age of participants varied from 7.8 years to 13.4 years (M=9.7 years, SD 1.2 years), and the program period varied from 5 to 10 months (M=7 months) to complete the 60–65 lessons. Such variation in lesson frequency was not ideal, but reflects the reality of school timetabling. There were 15 drop-outs whose scores were not included.

Pre-testing and post-testing were performed largely by me, with some individual testing performed by postgraduate students who had been trained in the administration of the chosen tests. The pre-tests and post-tests for both groups were seven months apart.

The program presenters were qualified primary reading teachers who had received at least minimal training in presenting the Decoding program.

Measures

A wide variety of tasks have been used to measure the construct of phonemic awareness. The Test of Phonological Awareness (TOPA) (Torgesen & Bryant,

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

1994) measures phoneme segmentation, one of the most relevant phonological awareness tasks to reading (Nation & Hulme, 1997). The test manual notes that the TOPA meets the requirements for technical adequacy according to standards of the American Psychological Association (1985, cited in Torgesen & Bryant, 1994).

Construct: Phonological Recoding in Lexical Access

Many studies have noted the higher error rate, and slower naming speed of disabled readers confronted with continuous lists of numbers, letters, pictured objects, and colours (Share, 1995). The difficulty is independent of semantic abilities—remaining evident when skilled and less skilled readers are matched on receptive vocabulary (Jorm, Share, Matthews, & Maclean, 1986). Nor does it appear that the speed and error rates are due to visual perceptual processes, but rather to greater difficulty in establishing phonological representations (Share, 1995). The theoretical link between naming tasks and reading involves the requirement of retrieving the name for a stimulus presented in visual format. In practice, it has been the speed with which the task is completed that correlates most highly with both word recognition and comprehension (Wolf, 1991).

A continuous picture naming test was developed for this study to provide a simple test of rapid naming—one directly relevant to reading. The skill has been assessed in a number of forms, but usually involves naming of known items: letters, numbers, colours, pictures, and objects. This test is a variant of the Rapid Automatised Naming test (Denckla & Rudel, 1976). It has been argued that letter naming is the naming skill most salient to reading, which is unsurprising given that it directly involves an element of the reading process, and is accepted as a strong predictor of future reading success in beginning readers. However, it was not assumed that all students were firm in their letter–sound knowledge, and likely that a number of the students would fall into this category. Additionally, using picture naming rather than letter naming avoids any reciprocal effects of reading ability upon letter naming (Johnston & Kirby, 2006).

The Picture Naming Test in this study uses black and white line drawings of everyday objects and events. The test comprised 60 pictures in 3 pages, and students were allowed one minute to name as many as they could.

Reliability figures (Hempenstall, 1995) were obtained by using a test–retest protocol with an interval of 2 weeks, involving a class of 28 students from a primary school involved in the study. The composite Year 3–4 class was tested individually in the identical format to the subsequent study. The ages of students ranged from 7.07 to 10.2 years. The Pearson correlation was .77.

Construct: Phonological Recoding in Working Memory

Working memory may not be a major limiting factor in skilled reading because most words are recognised instantly, and comprehension occurs at the time of the word's

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

fixation (Crowder & Wagner, 1992). For unskilled and novice readers, however, shortcomings in verbal working memory are likely to be exposed in the blending task, and in retaining the meaning of a sentence during its progressive decoding (Share, 1995). Disabled readers typically struggle to retain in working memory verbal material presented orally or visually (see Wagner & Torgesen [1987] for a review). Such short-term memory problems for verbal material have been evidenced in a variety of memory tasks including digits, letters, groups of words or sentences, and in objects and nameable pictures (Share & Stanovich, 1995). The performance of these tasks requires the capacity to store information represented in a phonological code. The deficit appears specific to phonological representation, as in visuo-spatial tasks there is no similar deficit (Share, 1995).

Thus, the relationship between memory span and reading is well established correlationally, but there is little evidence to support a direct causal role from memory to reading. Hulme and Roodenrys (1995) provide data to support the idea that short-term memory is merely a marker for other phonological deficits (especially, the quality of phonological representations).

Further, short-term memory impairment has been noted prior to school commencement, and hence cannot be explained as merely a consequence of slow reading progress; although interestingly, the ability may be amenable to improvement as reading skill develops (Ellis, 1990 ; Goldstein, 1976, cited in Share, 1995). Pre- and post-testing of Digit Span may detect any such effects occurring during the intervention.

The measure chosen for phonological recoding in working memory was the Digit Span subtest of the Wechsler Intelligence Scale for Children—Third Edition (Wechsler, 1991). This subtest is recognised as having well-established reliability and validity (Sattler, 1992).

Construct: Decoding

The Woodcock Reading Mastery Tests—Revised (Woodcock, 1987) is a comprehensive reading assessment tool frequently used in educational settings. The Word Attack subtest requires the student to decipher nonsense words. A correct response precludes the possibility of having used other than a phonological recoding strategy, or reading by analogy with similar real words. Validity and reliability are well regarded (Olson, Forsberg, Wise, & Rack, 1994).

This subtest has been used in a number of studies to assess phonological recoding (e.g., Alexander, Anderson, Heilman, Voeller, & Torgesen, 1991; Bowers, 1995; Bowers & Swanson, 1991; Bowey, Cain, & Ryan, 1992). The test is used here because it measures the degree to which students transfer phonemic awareness to the reading task. It also correlates strongly with word recognition and reading comprehension (Elbro, Nielsen, & Petersen, 1994; Vellutino, Scanlon, & Tanzman, 1994), and thus can arguably provide a proxy for general reading progress.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Construct: Spelling Ability

Spelling has been considered at least partly a phonologically-based skill (Plaza, 2003), and its correlation with phonemic awareness has been reported as about .6 (Shankweiler, Lundquist, Dreyer, & Dickinson, 1996; Stage & Wagner, 1992). Westwood (2005) also reports a correlation between spelling ability and reading achievement of around .89 to .92 from about age 8 years.

There have been a number of approaches used to assess spelling. A dictated word list approach was adopted because students are familiar with such a format, for ease of assessment in a group setting, and because it is a generally accepted format in educational research (Moats, 1994). The Brigance Comprehensive Inventory of Basic Skills (Brigance, 1992) spelling subtest is primarily a criterion-referenced instrument of this type. It is based on words used at the various grade levels in five or more of nine published spelling programs. It contains 10 words per year-level.

The test has limitations. For example, there have been no published reliability figures. Test–retest reliability was determined (Hempenstall, 1995) in a class of 28 students in one of the primary schools involved in the study. The composite Year 3–4 class was tested in a group format, using blank sheets of paper to cover their work in order to preclude collaboration. The ages of students ranged from 7.07 to 10.2 years. Pearson correlation was calculated at .97 (Statistical Package for the Social Sciences, Version 6.1, 1995).

The Corrective Reading Program

The *Corrective Reading* program is a remedial reading program designed for students in late Year 3 and above. It comprises two strands: Decoding and Comprehension, and within these strands are a number of levels. The Decoding strand was the focus of this study; the four levels (A, B1, B2, C) correspond to the students' decoding capacity, as assessed with a placement test.

The *Corrective Reading* program has been evaluated on many occasions, though its effects on phonological processes have not previously been a focus. Most analyses have emphasised word recognition and reading comprehension as outcome variables, and results for a wide range of poor readers have been strong (Grossen, 1998).

Selection

The placement test is administered prior to the program and consists of several passages of prose, the rate and accuracy of reading determining the program level for any given student. The placement test also ensures that student groups are relatively homogeneous in their decoding ability, and that they are neither over-challenged by the level of difficulty of the program, nor already competent at that level. The wait-list group provided the source of the non-equivalent control group.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Program Design

There are two major features evident in the *Corrective Reading* program. They are the emphasis on decoding skills (phonics) and the Direct Instruction approach to teaching the phonics content. It includes work on both isolated words and connected sentences, but its major emphasis is at the level of word structure. It is made clear to students that the decoding of novel words involves careful word analysis rather than partial cue or contextual guessing. Students are continually prompted to take account of all letters in a word, and become sensitised to common (and often problematic) letter groupings, for example, those beginning with combinations *st*, *bl*, *sl*, *fl*, *pl*, *sw*, *cl*, *tr*, *dr*; or ending with *nt*, *nd*, *st*, *ts*, *mp*, *ps*, *cks*, *ls*, *ms*, *th*, *er*, *ing*, *ers*, *y*. The sentences provided are constructed in a manner that allows few clues for contextual guessing, but provides ample opportunities to practise what has been learned in the teacher-presented word-attack segment of the lesson.

In this study, groups comprised about 10 students. Lessons are scripted, and use choral responses prompted by teacher signals. Teacher monitoring of responses helps determine the amount of repetition deemed necessary for mastery. Lessons typically range from 45 minutes to one hour, dependent on teacher lesson pacing. Program design specifies an optimum schedule of five lessons each week. This level of intensity has been found important for students with reading problems, as they tend to have difficulty retaining new skills and knowledge. There is strong emphasis on massed practice for mastery, and spaced practice for retention. If the lesson frequency is too low, retention may be jeopardised—leading to a general progress deceleration (Torgesen, 2003).

The Level A program focuses attention on word structure through reviewing letter–sound correspondence, and regular rhyming, blending and segmenting activities. It relates these phonemic awareness activities to the written word by initially emphasising regularly spelled words decomposable by using these skills. When this phonic approach is accepted by students as a viable (even valuable) strategy, common irregular words are introduced. In the program authors' view, this sequence reduces the jettisoning of the generative decoding strategies that may occur when irregular words are initially encountered at the high rate common in authentic literature.

Engelmann, Hanner, and Johnson (1999) describe the range of skills taught in Decoding A: letter/sound identification; sounding-out (segmenting) orally presented words, and then *saying them fast* (blending); decoding words of varying degrees of irregularity; reading whole words *the fast way*; reading short groups of words; sentence reading; spelling. Related skills such as matching letters, and common letter groupings (such as *ing*), word completion (for example, rhyming), and symbol scanning are included on the student worksheets. The sentence-reading exercises provide practice in reading words within a context. They are designed to retrain students in how to read words in sentences; achieved partly through ensuring contextual strategies will be unproductive, and through immediate correction of all decoding errors.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

The next level of the *Corrective Reading* program builds on the curriculum presented in Level A. The typical Decoding B lesson is divided into four major parts. In word-attack skills, students practise pronouncing words, identifying the sounds of letters or letter combinations, and reading isolated words composed of sounds and sound combinations that have been learned by the students.

In group story-reading, students take turns reading aloud from their storybook, while those who are not reading follow along. The stories are divided into parts, and when the group reads a story part within the error limit, the teacher presents specified comprehension questions for that part.

In individual reading checkouts, assigned pairs of students read two passages, the first of which is from the lesson just read by the group, whereas the second is from the preceding lesson. Each member of the pair first reads the passage from the current story, then a timed passage from the preceding lesson. Points for this passage are earned if the student reads it within a specified rate and error criterion.

Workbook activities conclude the lesson. The tasks are integrated with the activities in the other sections to provide additional practice opportunities.

In this study, 85 students participated in Level A and 49 in Level B of the program.

Results

Descriptive Statistics

Tables 1 and 2 provide the raw and transformed data used for all analyses.

Reading Disability Criterion

It has been argued that the major deficit facing the disabled reader is a difficulty in decoding single words, and that the primary basis for this difficulty is phonological in nature (Torgesen & Hudson, 2006). It has also been recognised that a pseudoword decoding test is an appropriate tool for discerning such a difficulty (Elbro, Nielsen, & Petersen, 1994; Hoover & Gough, 1990; Share & Stanovich, 1995; Stanovich, 1988). On the Word Attack sub-test of the Woodcock Tests of Reading Mastery (1987), the average score of the combined cohort is at the 5th percentile, an average delay of 2.8 years. This represents a score between 1.5 and 2 standard deviations from the mean, sufficient in most definitions for a diagnosis of reading disability (Felton, 1992; Lovett & Steinbach, 1997; Lovett et al., 1994; Lyon & Moats, 1997; Newby, Recht, & Caldwell, 1993; Prior, Sanson, Smart, & Oberklaid, 1995; Stanovich & Siegel, 1994; Vellutino et al., 1996).

Multivariate Analyses

A single-factor between-subjects multivariate analysis of covariance (mancova) was performed to indicate whether there was any difference between the experimental

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Table 1. Experimental vs. control group: mean raw scores

	n	Age	TOPA	Word Attack	PNT	Digit Span	Spelling
Control	72						
Pretest							
Range		83–153	5–20	0–29	6–54	4–17	5–39
M		115.72	14.57	11.38	33.43	9.28	17.85
SD		14.77	4.49	6.77	9.03	2.22	6.46
Posttest							
Range		90–160	6–20	0–28	14–52	5–15	4–34
M		122.61	15.28	12.31	34.57	9.57	19.53
SD		14.76	4.36	6.96	8.57	1.84	6.33
Experimental	134						
Pretest							
Range		92–161	1–20	0–31	12–53	3–15	1–31
M		115.27	12.84	10.16	33.03	9.34	15.47
SD		13.26	3.90	6.45	7.64	2.08	5.58
Posttest							
Range		97–167	6–20	0–41	9–55	2–18	7–41
M		122.24	17.04	19.54	37.39	10.31	20.99
SD		12.83	3.47	8.18	8.01	2.28	5.75

and control groups on the combined post-test scores for the five main dependent measures. The five corresponding pre-test scores served as covariates. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(25, 707.32)=2.33$, $p<.001$, so subsequent analysis required fitting separate slopes for

Table 2. Experimental vs. control group: mean power transformed scores

	n	Word Attack	Digit Span	Spelling
Control	72			
Pretest				
Minimum		1.00	0.06	2.72
Maximum		2.49	0.26	9.76
M		1.90	0.12	5.91
SD		0.31	0.03	1.37
Posttest				
Minimum		1.00	0.07	2.37
Maximum		2.47	0.21	8.97
M		1.94	0.12	6.26
SD		0.30	0.02	1.33
Experimental	134			
Pretest				
Minimum		1.00	0.07	1.00
Maximum		2.53	0.35	8.47
M		1.83	0.12	5.40
SD		0.34	0.03	1.28
Posttest				
Minimum		1.00	0.06	3.35
Maximum		2.72	0.18	10.07
M		2.20	0.11	6.58
SD		0.30	0.02	1.12

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

each level of the treatment group factor. This analysis revealed that there was a significant multivariate relationship between the combined pre-test scores and the combined post-test scores for both the experimental group, Wilks' λ=.19, F(25, 707.32)=16.13, p<.001, and the control group, Wilks' λ=.16, F(25, 707.32)=18.08, p<.001. However, with the pre-test results partialled out separately for the two groups, there was a significant overall difference favouring the treatment over the control group, Wilks' λ=.89, F(5, 190)=4.75, p<.001.

Results for the combined variables were also analysed using a two-way mixed multivariate analysis of variance (manova). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for group, Wilks' λ=.94, F(5, 200)=2.59, p=.027, power=0.79, and for time, Wilks' λ=.40, F(5, 200)=60.55, p<.001, power=1.00, and for the group-by-time interaction, Wilks' λ=.60, F (5, 200)=26.85, p<.001, power=1.00. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pre-test, Wilks' λ=.94, F (5, 200)=2.61, p=.026, multivariate effect size=.06, power=.80 and at post-test, Wilks' λ=.84, F (5, 200)=7.54, p<.001, multivariate effect size=.16, power=1.00. Further, a significant pre- to post-test difference was found for the control group, Wilks' λ=.72, F (5, 67)=5.22, p<.001, multivariate effect size=.28, power=.98, and for the experimental group, Wilks' λ=.22, F (5, 129)=93.78, p<.001, multivariate effect size=.78, power=1.00, and the magnitude of effect was substantially larger for the experimental group. The multivariate effect size ($1-\lambda$) can be considered large when it exceeds 0.15 (Cohen, 1988).

Univariate Analyses

This series of outcomes involved univariate analyses of the pre-test and post-test data, and also included the effect size *d*. Under the Cohen (1988) convention, 0.2 constitutes a small effect size, 0.5 a medium effect size, and 0.8 a large effect size. Slavin (1990) argued that an effect size above 0.25 should be considered educationally significant. Assumptions of normality and homogeneity of variance were tested for all data used in analyses of covariance and analyses of variance, and data transformations were performed when necessary.

Analyses were performed on the total sample of 206 students. The overall finding was that educationally significant change occurred in each of the measured variables, the size of the program effect varying from medium in the case of Digit Span, and Picture Naming, to large in Word Attack, TOPA, and Spelling.

Test of Phonological Awareness (TOPA)

Results for TOPA were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pre-test scores serving as the covariate and post-test

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 202)=14.15$, $p<.001$, so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pre-test scores covaried significantly with post-test scores for both the control, $F(1, 202)=127.84$, $p<.001$, and experimental groups, $F(1, 202)=57.69$, $p<.001$. With the pre-test results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 202)=31.73$, $p<.001$.

Results for TOPA were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204)=0.00$, $p=.98$, but a significant main effect was found for time, $F(1, 204)=172.29$, $p<.001$, power=1.00, and the group-by-time interaction, $F(1, 204)=53.75$, $p<.001$, power=1.00, which is illustrated in Figure 1. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pre-test, $F(1, 204)=8.23$, $p=.005$, $d=-0.48$, and at post-test, $F(1, 204)=10.04$, $p=.002$, power=1.00, $d=0.53$. Further, no significant pre- to post-test difference was found for the control, $F(1, 204)=3.41$, $p=.066$, $d=0.18$, power=0.451, but a significant pre- to post-test difference was found for the experimental groups, $F(1, 204)=222.63$, $p<.001$, $d=1.29$, power=1.00, and the magnitude of effect was large for the experimental group.

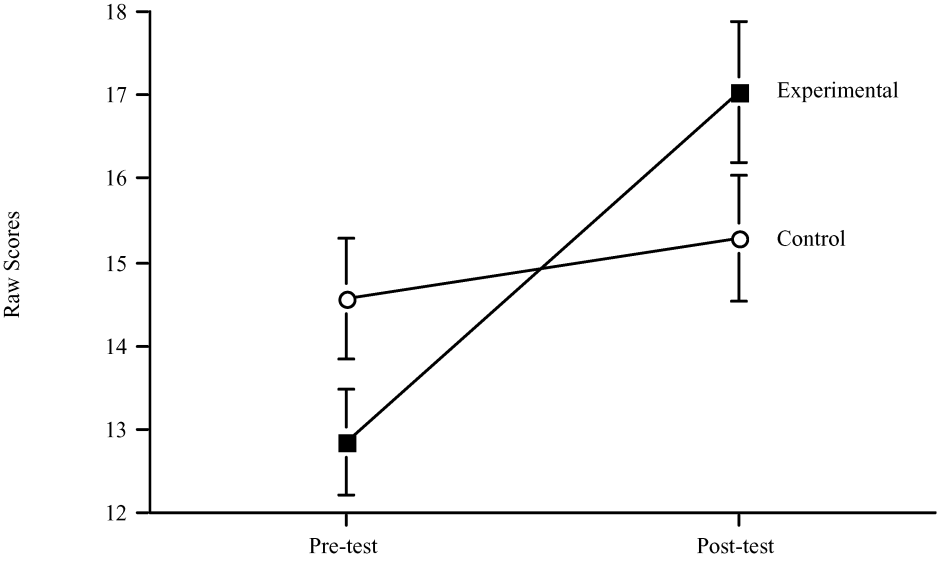


Figure 1. Interaction (+Standard Error) between experimental and control groups at pre- and post-test for TOPA

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Word Attack

Results for Word Attack were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pre-test scores serving as the covariate and transformed post-test scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 202)=11.28$, $p=.001$, so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pre-test scores covaried significantly with post-test scores for both the control, $F(1, 202)=101.96$, $p<.001$, and experimental groups, $F(1, 202)=85.88$, $p<.001$. With the pre-test results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 202)=23.55$, $p<.001$.

Results for the power transformed scores for Word Attack were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for group, $F(1, 204)=4.79$, $p=.030$, power=0.58, and for time, $F(1, 204)=196.06$, $p<.001$, power=1.00, and the group-by-time interaction, $F(1, 204)=73.49$, $p<.001$, power=1.00, which is illustrated in Figure 2. Follow-up testing of the interaction using simple main effects found a non-significant difference between the experimental and control groups at pre-test, $F(1, 204)=2.01$, $p=.158$, $d=-0.20$, power=.29, but a significant difference at post-test, $F(1, 204)=33.03$, $p<.001$, power=1.00, $d=1.00$. Further, no significant pre- to post-test difference was found for the control, $F(1, 204)=1.86$, $p=.174$, power=.27, $d=0.15$, but a significant pre- to post-test difference was found for the experimental groups,

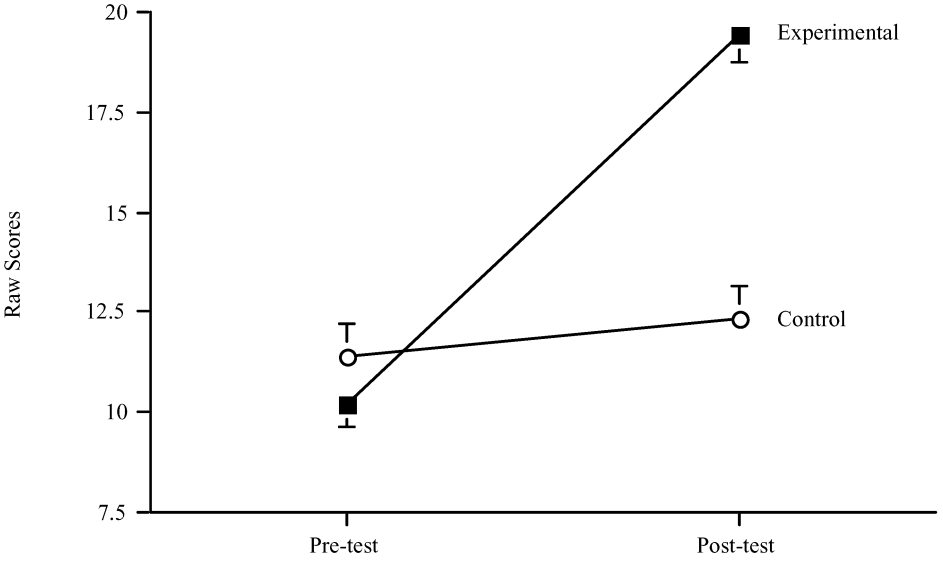


Figure 2. Interaction (+ SE) between experimental and control group at pre- and post-test for Word Attack

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

$F(1, 204)=267.69$, $p<.001$, power=1.00, $d=1.34$., and the magnitude of effect was large for the experimental group.

Picture Naming Test

Results for the Picture Naming Test were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pre-test scores serving as the covariate and post-test scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 202)=2.27$, $p=.134$. With the pre-test results partialled out there was a significant overall difference between the experimental and control groups $F(1, 203)=10.48$, $p=.001$.

Results for the Picture Naming Test were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204)=0.92$, $p=.337$, power=0.17, but a significant main effect was found for time, $F(1, 204)=47.49$, $p<.001$, power=1.00, and the group-by-time interaction, $F(1, 204)=10.11$, $p=.002$, power=.88, which is illustrated in Figure 3. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pre-test, $F(1, 204)=0.11$, $p=.737$, power=1.00, $d=-0.06$, but a significant difference at post-test, $F(1, 204)=4.22$, $p=.041$, power=.53, $d=0.39$. Further, no significant pre- to post-test difference was found for the control group, $F(1, 204)=2.28$, $p=.133$, power=.32, $d=0.15$, but a significant pre- to post-test difference was found for the

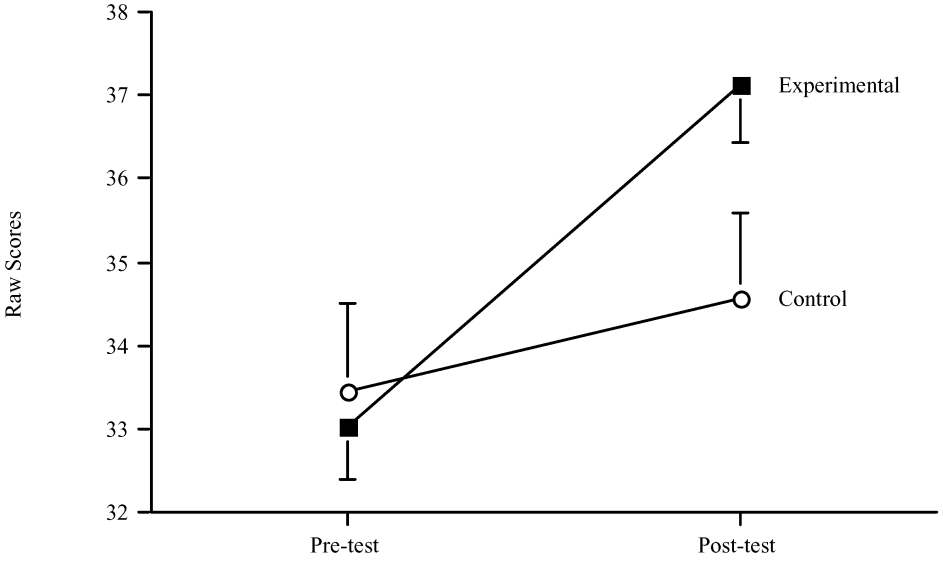


Figure 3. Interaction (+ SE) between experimental and control group at pre- and post-test for Picture Naming Test

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

experimental group, $F(1, 204)=55.31$, $p<.001$, power=1.00, $d=0.57$, and the magnitude of effect was medium for the experimental group.

Digit Span

Results for Digit Span were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pre-test scores serving as the covariate and transformed post-test scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 202)=0.25$, $p=.621$. With the pre-test results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 203)=7.92$, $p=.005$.

Results for power transformed scores for Digit Span were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204)=1.5$, $p=.222$, power=.23, but a significant main effect was found for time, $F(1, 204)=28.71$, $p<.001$, power=1.00, and not for the group-by-time interaction, $F(1, 204)=3.68$, $p=.056$, power=.48, which is illustrated in Figure 4. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pre-test, $F(1, 204)=0.00$, $p=.947$, power=.03, $d=0.03$, but found a significant difference at post-test, $F(1, 204)=6.08$, $p=.015$, power=0.69, $d=0.38$. Further, no significant pre- to post-test difference was found for the control, $F(1, 204)=2.62$, $p=.107$, power=.36, $d=0.16$, but a significant difference was found

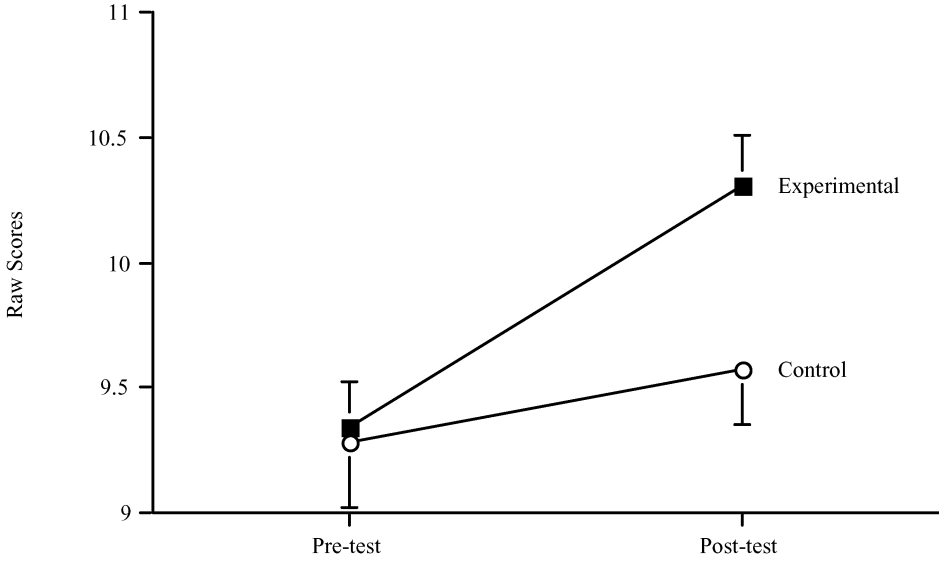


Figure 4. Interaction (+ SE) between experimental and control group at pre- and post-test for Digit Span

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

for the experimental group, $F(1, 204)=29.77$, $p<.001$, power=1.00, $d=0.48$, with a medium effect size for the experimental group.

Brigance Spelling

Results for Brigance Spelling were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pre-test scores serving as the covariate and transformed post-test scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 202)=5.37$, $p=.021$, so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pre-test scores covaried significantly with post-test scores for both the control, $F(1, 202)=126.58$, $p<.001$, and experimental groups, $F(1, 202)=112.42$, $p<.001$. With the pre-test results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 202)=12.26$, $p=.001$.

Results for the power transformed Spelling scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204)=0.30$, $p=.58$, power=.038, but a significant main effect was found for time, $F(1, 204)=188.89$, $p<.001$, power=1.00, and the group-by-time interaction, $F(1, 204)=36.89$, $p<.001$, power=1.00, which is illustrated in Figure 5. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pre-test, $F(1, 204)=7.03$, $p=.009$, power=.75, $d=-0.42$, but

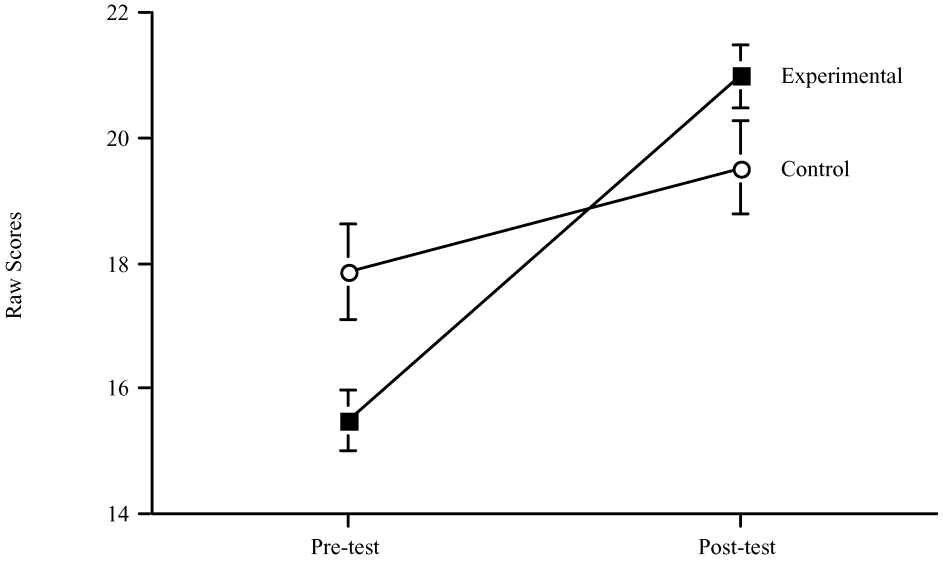


Figure 5. Interaction (+ SE) between experimental and control group at pre- and post-test for Brigance Spelling

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

not at post-test, $F(1, 204)=3.32$, $p=.07$, power=.44, $d=0.25$. Further, significant pre- to post-test differences were found for both the control, $F(1, 204)=10.41$, $p=.001$, power=.89, $d=0.27$, and experimental groups, $F(1, 204)=215.38$, $p<.001$, power=1.00, $d=0.99$; however. the magnitude of effect was large for the experimental group.

Is Success in the *Corrective Reading* Program Predicted by Any of the Pre-test Scores?

In addition to investigating the relationship among the phonological processes, another issue of interest was the potential of pre-test scores to predict which students would make good progress, and which would not. In prediction of gains in Word Attack for the experimental group, Table 3 indicates that program membership was by far the strongest. Whilst Word Attack and Spelling pre-test scores were significant predictors, their combined contribution is less than 7%—small in comparison with that of Program (almost 30%).

The initial questions were: Did participation in the *Corrective Reading* program increase phonemic awareness, phonological recoding (word attack) skills, and other phonological processes (naming, working memory)? Did the *Corrective Reading* program effects generalise to spelling? The results presented in the above sets of analyses indicated a clear pattern of statistically and educationally significant increases represented in the post-test scores for the experimental group. The effects varied from large (TOPA, Word Attack, Spelling) to moderate (Digit Span and Picture Naming).

Discussion

In this study of 206 disabled readers from several Melbourne primary schools, the *Corrective Reading: Decoding* program was implemented for 134 students, while 72 students on a wait-list provided a control. The program has a systematic, explicit phonics emphasis, with attention to letter–sound correspondences, and to the

Table 3. Summary of stepwise regression analysis for program and pre-test scores predicting word attack gains scores ($n=206$)

Step	Variable	R	R^2	δR^2	sig F	β	*p*
1	Word Attack	.194	.038	.038	.005	−.094	.005
2	Spelling	.261	.068	.031	.001	−.096	.011
3	Program	.604	.365	.296	.000	.555	.000
Constant in the final equation							.040
Variables not in the final equation							
	TOPA						.060
	PNT						.346
	Digit Span						.506

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

phonemic awareness skills of segmenting and blending. Pre-test and post-test of phonological processes, word attack, and spelling indicated statistically significant and educationally important changes in all variables for the experimental group.

All the students had received reading instruction in their schools prior to participating in the *Corrective Reading* program. Their failure to make adequate progress can be construed as arising from individual weaknesses, or from a failure of the schools' reading programs to elicit appropriate progress, or from some combination of the two. The general model of reading in this study places word-level processes at the centre of reading disability, and phonological processes as the major underlying abilities causal to reading development (Ehri, 1995). The outcomes of the study indicate that these skills can be developed, even in students who have had prior opportunity, but have been unable to do so in the context of earlier instruction.

That these phonological processes develop simultaneously with advances in word attack suggests that such skills remain important even for older students. That the developmentally earlier (phonetic decoding) stage should not be ignored has been emphasised by Share (1995), Share and Stanovich (1995), and by Shankweiler, Lundquist, Dreyer, and Dickinson (1996). This finding conflicts with the popular view that any phonic emphasis should be discontinued before Year 3, replaced by a new emphasis on orthographic processing and/or comprehension strategies.

Regression analyses were performed on the experimental group at post-test to add information about the relationship between the variables, and to consider whether pre-test variables were predictive of outcome for the experimental group. In analysing word attack gains, it was clear that the presence or absence of the program was the most powerful predictor by far. The results indicate that discernible and educationally significant change in word attack becomes evident within a relatively short period of time, approximately 50 hours over 7 months. These changes in word attack do not appear to be reliant on high levels of pre-existing phonological skills. For example, low picture naming speed at entry was not predictive of poor progress. It is likely that the environmental contribution of a carefully structured phonics program has sufficient influence to overcome any resistance to progress that may be associated with low initial naming speed. Nor was pre-existing phonemic awareness predictive of gains. This finding is consistent with that of Hogan, Catts, and Little (2005) who noted the predictive ability of phonemic awareness on word attack at Grade 2 but not at Grade 4, because the two variables become so highly correlated by that time. Another interpretation in this current study is that phonemic awareness has a reciprocal relationship with decoding, a view supported by previous research (Adams, 1990; Bowey & Francis, 1991; Wagner & Torgesen, 1987). Thus, systematic instruction in decoding can also boost phonemic awareness skill, at least for older students. This finding was also extended to beginning readers by Share and Blum (2005).

Perfetti, Beck, Bell, and Hughes (1987) noted that when structured code emphasis teaching was not provided, then initial levels of variables such as naming speed were

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

predictive of reading progress. They also argued that, when effective, phonically-based teaching occurred, the former levels of such variables were no longer predictive of progress. In fact, the effects of the intervention were to increase the level of phonological skills in the areas of naming speed and phonological recoding in working memory in addition to that of phonemic awareness. Their findings, replicated here, are consistent with both the reciprocal causation view and with the pre-eminence of phonological representation. The most common interpretation of such findings is that emphasis on the structure of words increases the quality or accessibility of phonological representations, and such change is represented in improved performance on the phonological variables. If, as they relate to reading, naming and working memory are reflective of an underlying variable (representation), there may be little value in attempting to influence these two variables through direct training of them. In the case of phonological recoding in working memory, an improvement following reading gains was noted by Wadsworth, DeFries, Fulker, Olson, and Pennington (1995) in their study involving the genetic analysis of twins. For phonological recoding in lexical access, Deeney, Wolf, and Goldberg O'Rourke (2001) noted how emphases on phonology, automaticity, and fluency (as seen in the Decoding program) enhance the reading of those with naming speed deficits.

An interpretation of the strong effect size for spelling is that students may have begun to perceive some logical structure behind spelling, rather than viewing it as an arbitrary and capricious system. Perhaps, the emphasis on word structure, especially the importance of each letter and its position in a word, may lead to a process analogous to Share's (1995) assertion of a self-teaching mechanism in reading. Davidson and Jenkins (1994) view the relationship of phonemic awareness and spelling as bi-directional, and these results are supportive of at least one of these directions. Burt and Butterworth (1996) assert a direct effect from phonological skills to spelling through the mnemonic enhancement of working memory, and an indirect effect through the benefits to spelling of enforced attention to letter sequence. It may also be that improved segmenting (a result of clearer phonological representations?) allows for more accurate conversion to spellings of the sounds in words. Stage and Wagner (1992) asserted that older students make less use of phonological processes in spelling than do young students, instead relying more on orthographic representations. It may be that this latter assertion refers only to older, *skilled* readers, and hence is really an assertion about stage rather than age.

Study Limitations

The group contrast in this study was between two distinct interventions—the schools' regular English program and the reading program. The reading program could be considered more motivating, and the improvement may have been partly based upon novelty. Alternatively, it could be argued that being withdrawn from class for a remedial program may be deflating to student motivation. In any case,

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

studies such as by Branwhite (1983) that extend over periods of a year and more continue to display strong effects—making the novelty explanation unlikely.

Statistical regression is a threat to internal validity, and there were some minor pre-test differences in that the intervention group had slightly lower scores than the control group on Phonemic Awareness, Word Attack, and Spelling, though not on the other two variables. These differences were partialled out in the analysis, though there remains the possibility that some unknown variable could account for the larger post-test improvement of the intervention group. Whatever that variable might be, it did not influence the reading placement test that judged the two groups to be homogeneous with respect to their reading instruction needs. In those schools in which there were both control and experimental groups the decision about which group received the treatment first was not based on problem severity. In other words, one would not expect regression toward the population mean to occur differentially across the groups.

As the experimental and control groups were in a variety of schools (State and Catholic) it seems unlikely that any extraneous events over the period of the program (historical threats to internal validity) could coincidentally affect only the experimental group.

Any effects on students of the test or testing procedure should have been equally distributed across both groups. These include student effects such as being sensitised by the pre-test, practice effects, and negative reactions to post-testing.

Issues of selection may jeopardise group comparability. For example, it is conceivable that schools prepared to provide a special reading program differ in important aspects from schools that are either unable to or choose not to do so. These school qualities may be efficacious in enhancing reading development but not obvious until the program's commencement, and the subsequent student progress falsely attributed to program effect. However, the control group comprised wait-list students, and was drawn from the same schools as those in the experimental group.

The students were markedly delayed in their literacy development (one to two standard deviations). Their improvement was significant, but they continue to require instruction in more advanced reading techniques, and in fluency and spelling. However, their learning trajectory was altered, and the risk of the further decline predicted by the Matthew Effects (Stanovich, 1986) was arguably somewhat diminished.

It is known that phonological process acuity is a strong predictor of reading success. What is not clear is whether these processes need to be directly addressed in order to assist reading development. The results of this study suggest that a focus on the task rather than the learner continues to be the best option for improving the achievement of those who currently struggle. Further, it points to the potential of systematic synthetic phonics programs to reduce the incidence of reading difficulties at the earliest instructional stage. The findings of this study are also consistent with the proposition by Torgesen, Wagner, Rashotte, Alexander, and Conway (1997) that remedial phonics programs for older students with a basic level of letter–sound

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

mastery and phonemic awareness (as were most students) may not require dedicated phonemic awareness programs.

It is apparent from research that early intervention (pre-school, Prep/Kindergarten, Year One) holds the greatest hope for reducing the deleterious effects of serious reading failure currently believed to impede up to 30% of all our students (Harrison, 2002; Livingstone, 2006; Louden, et al., 2000; Marks & Ainley, 1997).

However, there is probably another group (perhaps another 10–20%) whose progress becomes increasingly constrained by limited literacy the further they progress through secondary school. It is possible to enhance the prospects for both of these existing groups by intervening during their late primary and secondary schooling, and social justice requires us to provide for those students whom our system has failed. Intervention for these students is more difficult, but significant gains are achievable, and older students should not be ignored simply because early intervention is easier to implement and promote.

> So the sobering message here is that if children don't have the right experiences during these sensitive periods for the development of a variety of skills, including many cognitive and language capacities, that's a burden that those kids are going to carry; the sensitive period is over, and it's going to be harder for them. Their architecture is not as well developed in their brain as it would have been if they had had the right experiences during the sensitive period. That's the sobering message. But there's also a hopeful message there, which is unlike a critical period where it's too late. The sensitive period says: It's not too late to kind of try to remediate that later. And you can develop good, healthy, normal competencies in many areas, even if your earlier wiring was somewhat faulty. But it's harder. It costs more in energy costs to the brain. The brain has to work at adapting to earlier circuits that were not laid down the way they should have been. And from a society's point of view, it costs more in terms of more expensive programming, more specialized help. (Shonkoff, 2007, p. 13)

References

Adams, G., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.

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

Alexander, A. W., Anderson, H. G., Heilman, P. C., Voeller, K. K. S., & Torgesen, J. K. (1991). Phonological awareness training and remediation of analytic decoding deficits in a group of severe dyslexics. *Annals of Dyslexia*, *41*, 193–206.

American Federation of Teachers. (1997). *Building on the best: Learning from what works*. Retrieved November 11, 2003, from http://www.aft.org/Edissues/whatworks/index.htm.

American Institutes for Research. (1999). *An educators' guide to schoolwide reform*. Retrieved November 11, 2003, from www.aasa.org/Reform/index.htm.

American Institutes for Research. (2005). *CSRQ Center Report on Elementary School Comprehensive School Reform Models*. Retrieved February 12, 2005, from http://www.csrq.org/reports.asp.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

American Institutes for Research. (2006). *CSRQ Center Report on Elementary School Comprehensive School Reform Models (Updated)*. Retrieved December 11, 2006, from http://www.csrq.org/documents/CSRQCenterCombinedReport_Web11-03-06.pdf.

American Psychological Association. (1993, October). *Final Report of the Task Force on Promotion and Dissemination of Psychological Procedures*. New York: American Psychological Association, Division of Clinical Psychologists (Division 12).

Bentin, S., & Leshem, H. (1993). On the interaction between phonological awareness and reading acquisition: It's a two-way street. *Annals of Dyslexia*, *43*, 125–148.

Borman, G. (2007). *Taking reform to scale*. Wisconsin Center for Educational Research. Retrieved February 4, 2007, from http://www.wcer.wisc.edu/.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2002). *Comprehensive school reform and student achievement: A meta-analysis*. Report No.59. Washington, DC: Center for Research on the Education of Students Placed At Risk (CRESPAR), US Department of Education. Retrieved February 12, 2003, from http://www.csos.jhu.edu./crespar/techReports/report59.pdf.

Bowers, P. G. (1995). Tracing symbol naming speeds unique contributions to reading disabilities over time. *Reading and Writing*, *7*, 189–216.

Bowers, P. G., & Swanson, L. B. (1991). Naming speed deficits in reading disability. Multiple measures of a singular process. *Journal of Experimental Child Psychology*, *51*, 195–219.

Bowey, J. A., Cain, M. T., & Ryan, S. M. (1992). A reading-level design study of phonological skills underlying fourth-grade children's word reading difficulties. *Child Development*, *63*, 999–1011.

Bowey, J. A., & Francis, J. (1991). Phonological analysis as a function of age and exposure to reading instruction. *Applied Psycholinguistics*, *12*, 91–121.

Branwhite, A. B. (1983). Boosting reading skills by direct instruction. *British Journal of Educational Psychology*, *53*, 291–298.

Brigance, A. H. (1992). *Comprehensive Inventory of Basic Skills*. Melbourne, Australia: Hawker Brownlow.

Burt, J. S., & Butterworth, P. (1996). Spelling in adults: Orthographic transparency, learning new letter string and reading accuracy. *European Journal of Cognitive Psychology*, *8*(1), 3–43.

Carnine, D. (1995). Trustworthiness, usability, and accessibility of educational research. *Journal of Behavioral Education*, *5*, 251–258.

Castles, I. (1994). *SEIFA: Socio-economic indexes for areas*. Canberra, Australia: Australian Government Printer.

Coalition for Evidence-Based Policy. (2002). *Bringing evidence driven progress to education: A recommended strategy for the U.S. Department of Education*. Council for Excellence in Government's Coalition for Evidence-Based Policy. Retrieved December 12, 2002, from http://www.excelgov.org/displayContent.asp?Keyword=prppcEvidence.

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

Cook, T. D., & Campbell, D. T. (1979). *Quasi experimentation. Design and analysis issues for field settings*. Boston: Houghton-Mifflin.

Council for Exceptional Children. (1999). *Focussing on direct instruction*. Retrieved November 11, 2003, from http://dldcec.org/alerts/alerts_2.html.

Crowder, R., & Wagner, R. (1992). *The psychology of reading: An introduction*. New York: Oxford University Press.

Curriculum Review Panel. (2004). *Review of comprehensive programs*. Oregon Reading First Center. Retrieved January16, 2005, from http://reading.uoregon.edu/curricula/core_report_amended_3-04.pdf.

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

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Deeney, T., Wolf, M., & Goldberg O'Rourke, A. (2001). 'I like to take my own sweet time': Case study of a child with naming-speed deficits and reading disabilities. *Journal of Special Education*, *35*, 145–155.

Denckla, M. B., & Rudel, R. (1976). Rapid automatised naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*, 471–479.

Department for Education and Employment. (1998). *The National Literacy Strategy: Framework for teaching*. London: Standards and Effectiveness Unit.

Department of Education, Employment, and Training. (2001). *Successful interventions literacy research project*. Retrieved March 30, 2007, from http://www.eduweb.vic.gov.au/edulibrary/public/curricman/middleyear/research/successfulinterventions.doc.

Department of Education, Science, and Training. (2005). *Teaching reading: Report and recommendations*. National Inquiry into the Teaching of Reading. Australia: Melbourne.

Department of Education, Science, and Training. (2007). *Parents' attitudes to schooling*. Canberra, Australia: Australian Government. Retrieved June 4, 2007, from http://www.dest.gov.au/NR/rdonlyres/311AA3E6-412E-4FA4-AC01-541F37070529/16736/ParentsAttitudestoSchoolingreporMay073.rtf.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*(2), 116–125.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*, 250–287.

Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults. Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia*, *44*, 205–226.

Ellis, N. C. (1990). Reading, phonological processing and STM: Interactive tributaries of development. *Journal of Research in Reading*, *13*, 107–122.

Engelmann, S. (2004). The Dalmatian and its spots. *Editorial Projects in Education*, *23*(20), 34–35, 48. Retrieved February 3, 2004, from http://www.edweek.org/ew/ew_printstory.cfm?slug=20engelmann.h23.

Engelmann, S., Becker, W. C., Carnine, D., & Gersten, R. (1988). The Direct Instruction Follow Through Model: Design and outcomes. *Education and Treatment of Children*, *11*(4), 303–317.

Engelmann, S., & Bruner, E. C. (1984). *Reading Mastery fast cycle I, II*. Chicago: Science Research Associates.

Engelmann, S., & Bruner, E. C. (1988). *Reading Mastery*. Chicago: Science Research Associates.

Engelmann, S., Hanner, S., & Johnson, G. (1999). *Corrective Reading—Series Guide*. Columbus, OH: SRA/McGraw Hill.

Felton, R. H. (1992). Early identification of children at risk for reading disabilities. *Topics in Early Childhood Special Education*, *12*, 212–229.

Grossen, B. (Ed.) (1996). What was that project follow through? [Special issue] *Effective School Practices*, *15*(1), 1–85.

Grossen, B. (1998). *The research base for Corrective Reading, SRA*. Blacklick, OH: Science Research Associates.

Grossen, B. (2004). Success of a Direct Instruction model at a secondary level school with high-risk students. *Reading and Writing Quarterly*, *20*, 161–178.

Hammill, D. (2004). What we know about correlates of reading. *Exceptional Children*, *70*, 453–469.

Harris, R. E., Marchand-Martella, N. E., & Martella, R. C. (2000). Effects of a peer-delivered Corrective Reading program. *Journal of Behavioral Education*, *10*(1), 21–36.

Harrison, B. (2002, April). Do we have a literacy crisis? *Reading Reform Foundation*, *48*. Retrieved April 11, 2003, from http://www.rrf.org.uk/do%20we%20have%20a%20literacy%20crisis.htm.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

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

Hempenstall, K. (1995). *The Picture Naming Test*. Unpublished manuscript. Royal Melbourne Institute of Technology.

Hempenstall, K. (1996). The gulf between educational research and policy: The example of Direct Instruction and Whole Language. *Behavior Change*, *13*, 33–46.

Hempenstall, K. (2001). Some issues in phonics instruction. *EducationNews.Org*. Retrieved May 3, 2005, from http://www.ednews.org/articles/528/1/Some-issues-in-phonics-instruction-Implicit-and-explicit-phonics-instruction/Page1.html.

Hempenstall, K. (2003). The three-cueing system: Trojan horse? *Australian Journal of Learning Disabilities*, *8*(3), 15–23.

Hempenstall, K. (2006). What does evidence-based practice in education mean? *Australian Journal of Learning Disabilities*, *11*(2), 83–92.

Hoad, K-A., Munro, J., Pearn, C., Rowe, K. S., & Rowe, K. J. (2005). *Working Out What Works (WOWW) Training and Resource Manual: A teacher professional development program designed to support teachers to improve literacy and numeracy outcomes for students with learning difficulties in Years 4, 5 and 6*. Canberra, Australia: Commonwealth Department of Education, Science and Training; and Melbourne: Australian Council for Educational Research.

Hogan, T. P., Catts, H. W., & Little, T. D. (2005). The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools*, *36*(4), 285–293.

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

Hulme, C., & Roodenrys, S. (1995). Practitioner review: Verbal working memory development and its disorders. *Journal of Child Psychology and Psychiatry*, *36*, 373–398.

Johnston, R. S., & Watson, J. E. (2004). Accelerating the development of reading, spelling and phonemic awareness skills in initial readers. *Reading and Writing*, *17*, 327–357.

Johnston, T., & Kirby, J. (2006). The contribution of naming speed to the simple view of reading. *Reading and Writing*, *19*(4), 339–361.

Jorm, A. F., Share, D., Matthews, R., & Maclean, R. (1986). Behaviour problems in specific reading retarded and general reading backward children: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *27*(1), 33–43.

Livingstone, T. (2006, 14 October). Rating the system. *The Courier Mail*, p. 53.

Louden, W., Chan, L. K. S., Elkins, J., Greaves, D., House, H., & Milton, M. et al. (2000). *Mapping the territory—Primary students with learning difficulties: Literacy and numeracy*. (Vols. 1–3). Canberra, Australia: Department of Education, Training and Youth Affairs.

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

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

Lovett, M. W., & Steinbach, K. A. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children? *Learning Disability Quarterly*, *20*, 189–209.

Lyon, G. R. (1998). *Overview of reading and literacy initiatives. Statement to Committee on Labor and Human Resources*. Retrieved November 11, 2005, from http://www.nichd.nih.gov/publications/pubs/jeffords.htm.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Lyon, G. R., & Moats, L. C. (1997). Critical conceptual and methodological considerations in reading intervention research. *Journal of Learning Disabilities*, *30*, 578–588.

Malmgren, K. W., & Leone, P. E. (2000). Effects of a short-term auxiliary reading program on the reading skills of incarcerated youth. *Education & Treatment of Children*, *23*(3), 239–247.

Marks, G. N., & Ainley, J. (1997). *Reading comprehension and numeracy among junior secondary school students in Australia*. Melbourne: Australian Council for Educational Research.

McCluskey, N. (2003). *Best bets: Education curricula that work*. The Center for Education Reform. Retrieved July 18, 2003, from http://www.edreform.com/pubs/bestbets.pdf.

Moats, L. C. (1994). Assessment of spelling in learning disabilities research. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 333–350). Baltimore, MD: P. H. Brookes.

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly*, *32*(2), 154–167.

National Reading Panel. (2000). *Teaching children to read*. Retrieved November 11, 2005, from http://www.nationalreadingpanel.org.

Newby, R. F., Recht, D. R., & Caldwell, J. (1993). Validation of a clinical method for the diagnosis of two subtypes of dyslexia. *Journal of Psychoeducational Assessment*, *11*, 72–83.

Office of Management and Budget. (2007). *Program assessment: Reading First State Grants*. Retrieved May 11, from http://www.whitehouse.gov/omb/expectmore/summary/10003321.2006.html.

Office of the Victorian Auditor General. (2003). *Improving literacy standards in government schools*. Retrieved October 10, 2004, from http://www.audit.vic.gov.au/reports_par/Literacy_Report.pdf.

Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243–278). Baltimore, MD: P.H. Brookes.

Oregon Reading First. (2004). *Consumer's guide to evaluating supplemental and intervention reading programs Grades K–3: A critical elements analysis*. Retrieved January 11, 2006, from http://oregonreadingfirst.uoregon.edu/downloads/corrective_rdg_levela.pdf.

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly*, *33*, 283–319.

Plaza, M. (2003). The role of naming speed, phonological processing and morphological/syntactic skill in the reading and spelling performance of second-grade children. *Current Psychology Letters*. Special Issue on Language Disorders and Reading Acquisition. Retrieved March 27, 2007, from http://cpl.revues.org/document88.html.

Primary National Strategy. (2006). *Primary framework for literacy and mathematics*. London: Department of Education and Skills. Retrieved 26 October, 2006, from http://www.standards.dfes.gov.uk/primaryframeworks/.

Prior, M., Sanson, A., Smart, D., & Oberklaid, F. (1995). Reading disability in an Australian community sample. *Australian Journal of Psychology*, *47*(1), 32–37.

Purdie, N., & Ellis, L. (2005). *A review of the empirical evidence identifying effective interventions and teaching practices for students with learning difficulties in Years 4, 5 and 6*. A report prepared for the Australian Government Department of Education, Science and Training Camberwell, Victoria, Australia: Australian Council for Educational Research.

Reynolds, M., & Wheldall, K. (2007). Reading Recovery 20 years down the track: Looking forward, looking back. *International Journal of Disability, Development and Education*, *54*(2), 199–223.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Rosenshine, B. (2002). Helping students from low-income homes read at grade level. *Journal for Students Placed at Risk*, *7*. Retrieved March 19, 2003, from http://faculty.ed.uiuc.edu/rosenshi/Helping%20at-risk%20readers.htm.

Sackett, D., McRosenberg, W., Muir Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence-Based Medicine: What it is and what it isn't. *British Medical Journal*, *312*, 71–2. Retrieved November 11, 2005, from http://cebm.jr2.ox.ac.uk/ebmisisnt.html#coredef.

Sattler, J. M. (1992). *Assessment of children* (3rd ed.). San Diego, CA: Jerome M. Sattler.

Schacter, J. (1999). *Reading programs that work: A review of programs for pre-kindergarten to 4th grade*. Retrieved November 11, 2003, from http://www.mff.org/edtech/publication.taf?_function=detail&Content_uid1=279.

Shankweiler, D., Lundquist, E., Dreyer, L. G., & Dickinson, C. C. (1996). Reading and spelling difficulties in high school students: Causes and consequences. *Reading and Writing*, *8*, 267–294.

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

Share, D. L., & Blum, P. (2005). Syllable splitting in literate and preliterate Hebrew speakers: Onsets and rimes or bodies and codas? *Journal of Experimental Child Psychology*, *92*(2), 182–202.

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*, 1–57.

Shaughnessy, M. F. (2007). An interview with G. Reid Lyon: About Reading First. *EducationNews.Org*. Retrieved June 20, from http://www.ednews.org/articles/13053/1/An-Interview-with-G-Reid-Lyon-About-Reading-First/Page1.html.

Shonkoff, J. P. (2007). *The neuroscience of nurturing neurons*. Children of the Code. Retrieved March 1, 2007, from http://www.childrenofthecode.org/interviews/shonkoff.htm.

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, NJ: Prentice Hall.

Slavin, R. E. (2003). A reader's guide to scientifically based research. *Educational Leadership*, *60*(5), 12–16. Retrieved December 16, 2003, from http://www.ascd.org/publications/ed_lead/200302/slavin.html.

Slavin, R. E. (2004). Education research can and must address 'What Works' questions. *Educational Researcher*, *33*(1), 27–28.

Slavin, R. E. (2007). Statement of Robert E. Slavin, Director, Center for Data-Driven Reform in Education. Committee on Appropriations Subcommittee on Labor, Health and Human Services, Education, and Related Activities. Hearings on Implementation of No Child Left Behind. March 14, 2007. Retrieved March 16, 2007, from http://www.ednews.org/articles/8996/1/Statement-of-Robert-E-Slavin-Director-Center-for-Data-Driven-Reform-in-Education/Page1.html.

Smith, S. A. (2004). What is Corrective Reading? Tallahassee, FL: Florida Center for Reading Research. Retrieved November 11, 2005, from http://www.fcrr.org/FCRRReports/PDF/corrective_reading_final.pdf.

Snow, C. E., Burns, S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children. Report of the National Research Council*. Retrieved November 11, 2005, from http://www.nap.edu/readingroom/books/reading/.

Stage, S. A., & Wagner, R. K. (1992). Development of young children' phonological and orthographic knowledge as revealed by their spellings. *Developmental Psychology*, *28*, 287–296.

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

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities*, *21*, 590–604.

Stanovich, K. E. (1994). Constructivism in reading education. *Journal of Special Education*, *28*, 259–274.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, *86*(1), 24–53.

Statistical Package for the Social Sciences. [Computer Software] (1995). Chicago: SPSS.

Stone, J. E. (1996, April 23). Developmentalism: An obscure but pervasive restriction on educational improvement. *Education Policy Analysis Archives*. Retrieved November 11, 2005, from http://seamonkey.ed.asu.edu/epaa.

Texas, Office of the Governor. (1996). The Texas Reading Initiative: Mobilizing Resources for Literacy. Retrieved November 11, 2005, from http://www.governor.state.tx.us/_private/old/Reading/overview.html.

Torgesen, J. K. (2003). *Using science, energy, patience, consistency, and leadership to reduce the number of children left behind in reading*. Barksdale Reading Institute, Florida. Retrieved May 3, 2004, from http://www.fcrr.org/staffpresentations/Joe/NA/mississippi_03.ppt.

Torgesen, J. K., & Bryant, B. (1994). *Test of Phonological Awareness: Examiner's Manual*. Austin, TX: Pro-Ed.

Torgesen, J. K., & Hudson, R. F. (2006). Reading fluency: Critical issues for struggling readers. In S. J. Samuels, & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 130–158). Newark, DE: International Reading Association.

Torgesen, J., Wagner, R. K., Rashotte, C., Alexander, A., & Conway, T. (1997). Preventative and remedial interventions for children with severe reading disabilities. *Learning Disabilities*, *8*, 51–61.

Trezek, B. J., & Malmgren, K. W. (2005). The efficacy of utilizing a phonics treatment package with middle school deaf and hard-of-hearing students. *Journal of Deaf Studies and Deaf Education*, *3*, 257–271.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., & Chen, R. et al. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, *88*, 601–638.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1994). Component of reading ability: Issues and problems in operationalizing word identification, phonological coding, and orthographic coding. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 279–332). Baltimore, MD: P.H. Brookes.

Wadsworth, S., DeFries, J., Fulker, D., Olson, R., & Pennington, B. (1995). Reading performance and verbal short-term memory: A twin study of reciprocal causation. *Intelligence*, *20*, 145–167.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192–212.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.

Westwood, P. S. (2005). *Spelling: Approaches to teaching and assessment* (2nd ed.). Camberwell, Victoria, Australia: ACER Press.

White, W. A. T. (1988). A meta-analysis of the effects of direct instruction in special education. *Education & Treatment of Children*, *11*, 364–374.

Wolf, M. (1991). Naming speed and reading. The contribution of the cognitive neurosciences. *Reading Research Quarterly*, *26*, 123–141.

Woodcock, R. W. (1987). *Woodcock Tests of Reading Mastery—Revised*. Circle Pines, MN: American Guidance Service.

https://doi.org/10.1017/S1030011200025744 Published online by Cambridge University Press