Reading Recovery's unrecovered learners: Characteristics and issues

James W. Chapman★ and William E. Tunmer

College of Humanities and Social Sciences, Massey University, Palmerston North, New Zealand

Reading Recovery (RR) was developed in New Zealand in the early 1980s to provide 30 minutes of daily individualised literacy instruction over 20 weeks for students struggling with learning to read after one year of formal schooling. Considerable research has been undertaken on the RR programme. While results indicate short-term success for some students, each year 15–30% of students do not successfully complete the programme and are therefore 'unrecovered'. Research on the characteristics of these unrecovered students is sparse. This review examines findings on the characteristics of unrecovered students. These RR students typically have limited phonemic awareness and phonemically based decoding skills, and lower scores on RR screening measures on entry to RR than 'recovered' students. In New Zealand, unrecovered students tend to be enrolled in schools serving lower socio-economic neighbourhoods, and tend to be from Māori or Pasifika (Polynesian Pacific Island heritage) backgrounds. These students typically receive more RR lessons than recovered students. We conclude that RR does not tailor instruction to meet the needs of individual students, as claimed. The RR instructional model, developed in the 1970s, fails to recognise the importance of explicit, systematic instruction in phonemic awareness and the use of letter–sound relations. Such instruction is essential for most students who struggle with literacy learning during their early years of schooling and especially important for students who experience the most difficulty with learning to read. Suggestions are presented for strengthening the RR programme and for reducing the number of unrecovered students.

Introduction

Reading Recovery (RR) is said to be 'by far the most widely researched and used tutoring program in the world' (Slavin et al., 2011, p. 6). Developed in New Zealand by Marie Clay during the 1970s, RR was designed to help students identified as making limited progress in reading after 1 year of formal reading instruction (Clay, 1979). The programme was implemented across New Zealand during the 1980s (from 1983), and adopted in parts of Australia, England, Wales, Republic of Ireland, Northern Ireland, Denmark, Canada, Barbados and the USA during the 1980s and 1990s (McNaughton, 2007).

Research on the effectiveness of the RR programme is extensive with mixed and controversial results. Many reviews of RR conclude that the programme is effective and beneficial (e.g. D'Agostino & Murphy, 2004; US Department of Education, What Works Clearinghouse, 2007, 2013, 2014; D'Agostino & Harmey, 2016; D'Agostino et al., 2017; Shanahan, 2017). Other reviews, however, cast doubt on

---

★Corresponding author. College of Humanities & Social Sciences, Massey University, PN342, Private Bag 11 222, Palmerston North, New Zealand 4442. Email: j.chapman@massey.ac.nz

the efficacy of the programme and question both claims made in favour of the programme as well as the underlying instructional model (e.g. Wasik & Slavin, 1993; Hiebert, 1994; Shanahan & Barr, 1995; Grossen et al., 1999; Elbaum et al., 2000; Slavin et al., 2011; Chapman et al., 2015; Chapman & Tunmer, 2016).

Despite debates about many aspects of RR, including the extent to which students benefit and whether the benefits last beyond exit from the programme (e.g. Center et al., 1995; Chapman et al., 2001; Reynolds et al., 2009; Nicholas & Parkhill, 2013; Jesson & Limbrick, 2014), there is evidence that some students derive benefit from successful completion of RR. There is also evidence, however, that many students—between 10% and 30% of those who enter RR—do not gain from RR. Yet, surprisingly given the 35-year history of RR, little is known about the characteristics of students who are accepted into RR, but who do not successfully complete the programme and are therefore, unrecovered.

Understanding the characteristics of those unrecovered students is important for four reasons: (1) the claims made by Clay about the long-term effectiveness of the programme in substantially reducing reading failure; (2) the widespread use of the programme in numerous countries, referred to at the outset of this paper; (3) the considerable cost associated with RR compared with other early intervention programmes (e.g. Grossen et al., 1999; Hollands et al., 2013); and (4) the consistent rate of unsuccessful programme completions that has persisted since RR's inception.

The purpose of this review is to examine evidence regarding the efficacy of RR, with particular focus on the characteristics of unrecovered students, and to consider the reasons for students being unrecovered from RR. Following a summary outlining the development, nature and theoretical framework of RR, we present data relating to the characteristics of unrecovered students. Following that, we address factors strongly linked to students' outcomes following placement in RR. These factors include the location of students on the developmental continuum for word reading from pre-alphabetic to consolidated alphabetic phases, the effect of the nature of literacy instruction in the regular classroom and the effect of the amount of phonological skill training in RR. We conclude by suggesting ways in which the effectiveness of RR could be improved.

The goals of Reading Recovery

Based on Clay's observational research as a developmental psychologist during the 1960s, the RR programme was designed to cater for children whose reading progress was in the lowest 15–20% of the enrolment cohort in any given school (Clay, 2007). Specifically, Clay argued that RR was for the 'lowest achieving children' (1991, p. 60), those children who were 'hardest to teach' (2005a, p. i). She was adamant that the programme was to serve the lowest achievers in text reading, '*not excluding any child in regular six year old classrooms for any reason*' (Clay, 1991, p. 360; italics in the original).

Clay's confidence about the effectiveness of RR largely explains adoption of the programme throughout New Zealand and in other countries. Clay claimed that RR is a

programme which should clear out of the remedial education system all the children who do not learn to read for many event-produced reasons [i.e. environmental, cultural, or socio-economic causes] and all the children who have organically based problems but who can be taught to achieve independent status in reading and writing despite this . . . . (Clay, 1987, p. 169; words in brackets added)

Promotional statements about RR reflect Clay's confidence in the programme. The New Zealand RR website states:

The aim of Reading Recovery is to prevent literacy difficulties at an early stage before they begin to affect a child's educational progress. Providing extra assistance to the lowest achievers after one year at school, it operates as an effective prevention strategy against later literacy difficulties. Nationally, it may be characterized as an insurance against low literacy levels. (New Zealand Reading Recovery, 2018a, p. 1)

Similarly, University College London's Institute of Education in England, which hosts the RR programme in the UK, claims 'that no other system achieves such good results so swiftly and so long lasting' as RR (University College London, 2018, p. 1). The programme ensures that 'the lowest attaining children will succeed in literacy', and that RR 'will not just narrow the attainment gap for disadvantaged children, [it] will close it' (University College London, 2018, p. 1).

In the USA, the Reading Recovery Council of North America (RRCNA) states that the RR programme 'has a strong tradition of success with the lowest-achieving children' (Reading Recovery Council of North America, 2018, p. 1). Citing Wilson and Daviss (1994), the RRCNA maintains that RR 'is the best evidence yet of the direct link between good design and education excellence' (RRCNA, 2018, p. 1).

Many studies have been reported on RR. Given its 30-plus year history, surprisingly few studies have found robust evidence in support of the claims made by Clay and RR organisations in various countries. There are research results showing that RR meets the needs of some students (e.g. Pinnell et al., 1994; D'Agostino & Murphy, 2004; Schwartz, 2005; U.S. Department of Education, What Works Clearinghouse, 2007, 2013, 2014; May et al., 2015; D'Agostino & Harmey, 2016; D'Agostino et al., 2017; Shanahan, 2017). Many studies and reviews, however, cast considerable doubt on the effectiveness of the programme (e.g. Wasik & Slavin, 1993; Hiebert, 1994; Center et al., 1995; Shanahan & Barr, 1995; Grossen et al., 1999; Elbaum et al., 2000; Reynolds et al., 2009; Slavin et al., 2011; Nicholas & Parkhill, 2013; Chapman et al., 2015; Chapman & Tunmer, 2016).

Development and nature of Reading Recovery

Marie Clay developed RR during the 1970s (Clay, 1979) at the University of Auckland, where she was a member of the academic staff. Following results of field trials that were claimed by Clay (1979, 1980) to be successful, the New Zealand Department of Education funded the programme for adoption by schools throughout the country during the 1980s (Soler & Openshaw, 2006; Clay, 2007). RR was designed as a preventative early intervention programme. The focus of the programme is on the lowest achieving students who have been in school for 12 months, and who have not benefitted from formal reading instruction (Clay, 1985, 1993a, 2007).

The general aim of RR is to accelerate to average levels of reading performance the development of 6-year old students who show signs of reading difficulty during their first year of schooling. In this context, reading difficulty is operationally defined as reading progress that is in the lowest 15–20% of same-age peers in any given school that offers the programme. Progress in reading is assessed after 1 year of schooling with measures that constitute the Observation Survey of Early Literacy Achievement (Clay, 1993b). Included in this battery is a letter knowledge task, a sight word test, a measure of writing vocabulary, a dictation task, and a running record of the student's oral reading of a text.

Students selected for RR receive 30 minutes of daily individual instruction over an average of around 20 weeks (New Zealand Ministry of Education, 2016, p. 22) by specially trained RR teachers. Instruction is supposed to be customised to the individual needs of each student, with careful attention being given to changes in the student's progress (Clay, 1998). Special emphasis is placed on each student developing a 'self-extending system of reading strategies' (Clay, 2002, p. 33). This approach emphasises the ability to use in a flexible manner multiple cues (syntactic, semantic, visual, graphophonic) to identify unfamiliar words and correct errors while reading connected text (Clay, 1991).

Typical RR lessons consist of re-reading familiar books, reading the previous day's new book and taking a running record of student's reading, working with letters and words using magnetic letters, writing a story, putting together in the correct order a cut-up story, and reading a new book. Decisions regarding discontinuation (i.e. successful completion) from RR are based on the following criteria: reading at a level near the class average, becoming reasonably independent in reading and being in the programme for around 20 weeks.

Clay (2007) expected that 90% of students who entered RR would successfully complete the programme with the skills to read independently on a par with the average group of their classroom peers, without the need for further assistance. Roughly 10% of students who entered RR would not be successful and would require further ongoing assistance. The RR nomenclature for these students is 'referred on', or 'referred'. In this paper, we consider these students to be unrecovered. Clay (2007) expected that only around 1–2% of an age cohort would require ongoing assistance with literacy learning 'which starts early and continues for several years' (p. 16).

The RR lessons are designed to supplement the regular classroom reading programme. In New Zealand RR is compatible with the whole language approach to beginning literacy instruction (Smith & Elley, 1994; Chapman et al., 2015). However, in support of the expansion of the programme in other countries, Clay (1993a) maintained that RR works well in conjunction with all types of classroom literacy programmes. She stated that RR can be used with students 'from any kind of classroom programme . . . supplementary to the ongoing activities of the classroom' (p. i).

Reading Recovery theoretical framework

New Zealand has followed a strongly constructivist approach to literacy education for over 25 years. This approach is based on the assumption that learning to read is essentially like learning to speak, where both abilities are said to develop 'naturally'

(Smith & Elley, 1994, p. 81). As Wilkinson et al. (2000) noted, 'New Zealand's literacy practices have a long history of association with a developmental constructivist bias in teaching and learning' and 'direct instruction of specific knowledge and skills according to prespecified routines finds little favour' (p. 12).

Through the RR programme, Clay (1987, 1991, 1998, 2005a, 2005b) played a very strong role in the development of this constructivist approach. She helped to define the theoretical underpinnings for general literacy instruction as well as for the RR approach to early intervention. In the constructivist approach, there is little or no explicit, systematic teaching of phonemic awareness (the ability to reflect on and manipulate the phonemic elements of spoken words) and alphabetic coding skills (the ability to translate letters and letter patterns into phonological forms). Both phonemic awareness and alphabetic coding skills are necessary for learning to read successfully (e.g. Snow & Juel, 2005; Pressley, 2006; Tunmer et al., 2015).

RR is based on the multiple cues theory (or searchlights model) of reading (see Greaney, 2011). Skilled reading is viewed as a process involving minimal word-level information for identifying unfamiliar words in text. Instead, students are encouraged to make predictions about the identity of unknown words based on multiple sources of information. Clay (1991) stated:

> In efficient rapid word perception the reader relies mostly on the sentence and its meaning and some selected features of the forms of words. Awareness of sentence context (and often the general context of the text as a whole) and a glance at the word enables the reader to respond instantly. (p. 8)

Learning to read is seen largely as a process in which students learn to use multiple cues for figuring out unknown words in text. Text-based cues (picture cues, sentence-context cues, preceding passage context, prior knowledge activated by the text) are employed to generate predictions about what unknown words might be; letter-sound cues are used very sparingly and mainly to confirm language predictions. Clay (1998) argued that beginning readers

> Need to use their knowledge of how the world works; the possible meaning of the text; the sentence structure; the importance of order of ideas, or words, or of letters; the size of words or letters; special features of sound, shape and layout; special knowledge from past literacy experiences *before* they resort to left to right sounding out of chunks or letter clusters or, in the last resort, single letters. (p. 9, emphasis added)

In line with this view, explicit attention to the development of word analysis skills runs counter to the instructional philosophy of RR. Clay (1993a) emphasised the importance of students developing the flexible use of multiple cues to identify and correct mistakes while reading (Clay, 2005a, 2005b). Struggling readers are taught to use various strategies involving predicting, cross-checking and confirming when trying to identify unknown words when reading text. She wrote that meaning is 'the most important source of information', and that 'the most important test for the child is "Does it make sense?"' (Clay, 1991, p. 292). Word-level information, according to Clay, is used primarily for confirming language predictions: 'The child checks language predictions by looking at some letters' and 'can hear the sounds in a word he speaks and checks whether the expected letters are there' (Clay, 1993a, p. 41). Clay

(1993a) specifically indicated that children should be taught not to rely too heavily on word-level cues: 'if the child has a bias towards letter detail the teacher's prompts will be directed towards the message and the language structure' (p. 42). This means that when students show a preference for using word-level information to identify unknown words in text, Clay recommended that the teacher should divert their attention away from such information.

The benefits of RR are variable; some students benefit, at least in the short term, while others do not benefit. This variable value can be explained by current theories of reading development. Such views consider that reading acquisition is a cognitive-development process that occurs over time and that may be disrupted at different points (Tunmer & Nicholson, 2011). Many struggling readers take longer than their peers to develop the self-improving alphabetic coding skills necessary for making progress in learning to read, largely because their knowledge of print when they start learning to read is limited, and/or because of a developmental delay in the acquisition of crucial phonological awareness skills. For these students, the focus on text reading in RR lessons provides them with further practice in applying their developing alphabetic coding skills to word identification. As their reading efforts become increasingly successful, these delayed readers start making more independent use of letter-sound information to identify unknown words, from which further spelling–sound relations can be induced without explicit instruction. The extra practice in text reading provided in the RR programme, is therefore likely to be useful in helping some of those delayed readers catch up to their peers.

A large number of struggling readers, however—probably most students with literacy learning difficulties—perform at even lower developmental phases of word learning. Ehri (2005) described these as the pre-alphabetic and partial-alphabetic phases. Typical characteristics of delayed readers who are still in these phases of reading development include limited or severely limited phonemic and alphabetic coding skills. Students who have ongoing problems in identifying phonemic sequences in words will struggle to fully understand the alphabetic principle and learn spelling-to-sound relationships. More intensive and systematic instruction in phonemic awareness and phonemically based decoding skills is usually required for these students than what is normally provided in RR lessons.

Two studies support the claim that delayed readers who have not fully grasped the alphabetic principle and who have not had pre-school language experiences that facilitate the development of phonemic awareness benefit from more intensive, explicit and systematic instruction in phonemic awareness and phonemically based decoding skills than is typically provided in RR lessons. Iversen and Tunmer (1993) found in a US Rhode Island study that significant improvements in the effectiveness of RR could be achieved. Including more intensive and explicit instruction in phonological awareness and the use of letter–sound relations, in combination with strategy training on how and when to use this knowledge, was more beneficial in helping students to identify words while reading text and to spell words while writing messages than the standard RR approach.

Chapman et al. (2001) found in a New Zealand longitudinal study of RR that discontinued students who made only modest progress or who were unable to maintain gains made in the programme (most of the RR students in the study), had limited or

severely limited phonemic awareness and phonemically based decoding skills. Limitations in these important areas were observed prior to entry into RR, as well as during their first year in school, and during the year after they had been discontinued from the programme (Year 3). A small number of students obtained some modest benefit from RR. They had more advanced phonological skills prior to entry into RR than students who experienced little or no benefit from the programme. Moreover, progress in learning to read after taking part in RR was strongly related to phonological skills at the time of discontinuation. These results are consistent with the widely held view among reading scientists that the primary characteristic of developmental reading problems is poor context-free word recognition ability and associated phonological processing difficulties (Vellutino et al., 2004; Vellutino & Fletcher, 2005; Tunmer & Greaney, 2010; Tunmer et al., 2015).

These findings, together with the instructional model on which RR is based, support our view that RR is not tailored to meet the needs of individual students, as claimed by Clay (1998). Rather, RR persists with an instructional model that is essentially one-size-fits-all, suited mainly for middle class students who have relatively mild reading difficulties and sufficient language skills to benefit, at least in the short term, from placement in the strongly text-oriented RR programme. In that sense, RR appears to be of some benefit to those students who have mild reading difficulties, as shown in some of the studies that claim RR is a successful programme.

Limited evidence in support of Reading Recovery

Despite being developed initially for use in New Zealand, no robust research exists to show that RR is effective in that country (Chapman & Tunmer, 2011; Chapman et al., 2015). Clay's (1979) own research on RR that led to its introduction throughout New Zealand (and in other countries) was criticised because of serious design flaws (Shanahan, 1987; Nicholson, 1989; Iversen & Tunmer, 1993; Wasik & Slavin, 1993; Center et al., 1995). These flaws included no appropriate control group; inappropriate use of multiple *t*-tests for analysing gain scores; including only RR students who were discontinued rather than all RR students; and using only assessments developed by Clay rather than independent standardised measures.

Notwithstanding the shortcomings in her research, Soler and Openshaw (2006) reported that introduction of the programme throughout New Zealand occurred largely because politicians saw RR as a means of addressing concerns about the increasing incidence of reading difficulties and overall literacy standards. Similarly, politicians in England saw RR as a quick way to address literacy issues in that country. Following a visit to New Zealand in 1992 by two Inspectors from England to observe RR lessons and meet with RR teachers, the official report indicated that the programme had attained 'almost unmitigated success' (Soler & Openshaw, 2006, p. 105).

Numerous reports and reviews referred to earlier indicate that RR can be an effective intervention for some students, producing short-term gains in reading performance. Shanahan (2017), for example, summed up the evidence in support of RR as follows: 'its immediate impact on learning seems to be beyond question, given that it works under so many conditions' (p. 509).

While Shanahan's conclusion might have some validity in terms of short-term benefits, two New Zealand studies cast doubt on the longer term effectiveness of the programme. Data for 95 successfully discontinued RR students in Years 4 to 6 (ages 8 to 10) were analysed by Nicholas and Parkhill (2013). Almost 45% of these students scored in the stanine range of 1 to 3 (23rd percentile and below) on a New Zealand standardised test of reading comprehension (Progressive Achievement Tests, Darr et al., 2008). Scores for unrecovered students were not reported. Nicholas and Parkhill observed that initial, short-term gains made by many students as a result of RR were 'not sustained for almost half of the students' (p. 9).

In a similar study involving 342 discontinued students in Years 4 to 6, Jesson and Limbrick (2014) reported that close to 65% of students were performing below stanine 5 (40th percentile and below) on standardised measures of reading comprehension. In line with the Nicholas and Parkhill (2013) results, Jesson and Limbrick (2014) found that close to 40% were performing in the considerably below average range of stanines 1 to 3 (23rd percentile and below). These authors observed that short-term gains were achieved by many students as a result of RR. However, two to four years following successful completion of RR large numbers of students were not achieving at expected age-appropriate levels on standardised reading and writing tests, and their reading achievements were not at the average levels attained by their cohort peers in their schools. As with the Nicholas and Parkhill study, results were not reported for unrecovered students.

In light of these and other findings on RR, the question we address in this paper is whether the programme is an effective intervention for students who most need literacy supports, that is, the 'lowest achieving children' (Clay, 1991, p. 60). Prior to addressing this question, however, we examine evidence regarding selection practices for admitting students into the RR programme. Selection of students is an important issue that has an impact on the interpretation of claims about the programme's success.

Does Reading Recovery always select the lowest achieving children?

Research from published studies on the effectiveness of RR seldom addresses the extent to which the lowest achievers in a given cohort were actually placed in the programme. In New Zealand, those 6-year-olds with the most need for literacy supports do not always receive RR in schools that offer the programme (Church, 2005; Clay, 2005a; McDowall et al., 2005; Chapman et al., 2007) because of doubts about the programme's effectiveness for these students. McDowall et al. (2005) reported in a Ministry of Education-funded study of RR that teachers interviewed in the study mentioned the practice of not placing students in the programme who had the most need for literacy supports. They also mentioned the practice of withdrawing students from the programme if their improvement in reading was considered to be too slow.

In support of this view, Belgrave (2009) reported that RR teachers in her study told her that they manipulated which students they accepted into the programme, 'so as to have a degree of success with their students' (p. 51). Teachers commented to Belgrave, taking the 'very bottom students is a waste of time and money for all involved and that it is better to take on students that will benefit from the sessions' (p. 51).

Such restrictive practices extend beyond New Zealand. In an Australian study, Serry et al. (2014) also reported views of RR teachers questioning the effectiveness of RR for students who struggled the most with learning to read. Most of the RR teachers in their study believed that the RR 'intervention alone cannot be relied on to return a child to the *average* [original emphasis] reading level of his/her classroom' (Serry et al., 2014, p. 68). Serry et al. also found that RR teachers were not convinced that the programme could properly meet children's specific literacy learning needs, such as specific reading disability or dyslexia (p. 68). The authors further reported that RR teachers were essentially unable to properly describe the basic characteristics of specific reading disability and were ill-prepared to identify reading disability subtypes, despite claims by Clay that RR tailors teaching to the individual and specific needs of each student (Clay, 1998, 2005a, 2005b). Rather, RR teachers expressed feelings of being 'powerless' (Serry et al., 2014, p. 69) to meet the needs of those students who struggled the most with literacy learning. Well established underlying factors of reading disability, such as phonological processing and language difficulties, were not seen by RR teachers as critical factors that frequently impeded the progress of students in the programme. Serry et al. questioned whether the costly RR programme should continue to be delivered by schools given that there appears to be 'evidence that it may not be the best option for low progress readers' (p. 71).

In the USA RR also appears to restrict entry into the programme. A US$55 million 'scale-up' study funded mostly by the US Department of Education as part of the 2010 Obama administration economic stimulus package, involved more than 1400 schools with targeted assistance for over 88,000 students. The authors of the study (May et al., 2015) described a two-tier process for selecting students into the RR programme. They drew attention to the requirement of the RRCNA Standards and Guidelines (Reading Recovery Council of North America, 2009) that schools select 'the lowest achieving first-grade students'and serve the 'lowest scorers first' (May et al., 2015, p. 24). The May et al. study, however, did not consistently follow this selection requirement. Instead, schools initially identified a pool of 'low achieving students' (p. 24), not necessarily the lowest achieving students. From that pool, students were then selected for placement in the RR intervention. Some schools made significant departures from the RRCNA guidelines, by using various assessments (not necessarily only the assessments that constitute the Observation Survey) and a range of team approaches to identify which students should be placed in RR for this scale-up study. May et al. noted that 'in a number of schools, students were nominated to the pool based on classroom teachers' general observations or "gut" impressions of their needs' (p. 24).

The authors also commented on the 'extent to which schools excluded particular groups of students from receiving Reading Recovery' (p. 25). The reason given for excluding students was that preference would be given to place students in RR who were considered 'more likely to benefit from the intervention' (p. 25). These practices suggest that some (perhaps many) schools have little confidence in the effectiveness of RR in meeting the needs of the lowest achieving 6-year-old literacy learners.

Given this selective approach to the placement of students in the RR programme, the credibility of this scale-up study in terms of the goals of RR (meeting the needs of the 'lowest achievers') is questionable. Several other problems with the study render

the results inadequate for making the case in favour of widespread implementation in the USA (or elsewhere). These shortcomings include the fact that control students received a variety of different learning experiences, unsupported claims that RR led to ongoing literacy gains, very modest outcomes for students who successfully completed the programme, and significant issues with the assessments that were used (Chapman & Tunmer, 2016; Cook et al., 2017).

In the United Kingdom, an anecdotal report is consistent with findings about selection practices in Australia, New Zealand, and the USA. McLachlain (2014) wrote, 'the RR teacher in the UK does not select children for the scheme if they have any sort of learning difficulties and those who make no progress are discontinued from the scheme early and eliminated from the stats as if they never took part' (p. 39). As in New Zealand (Chapman et al., 2015) and Australia (Serry et al., 2014), McLachlain (2014) wrote that 'Most children who receive RR go on to need more intervention' (p. 39).

Clay was strongly opposed to such selection practices. She was resolute that the RR programme was for the 'hardest to teach children'regardless of the suspected cause of reading difficulties (1993a, p. i). She acknowledged, however, that such practices occurred: 'Schools have wanted to select children for the intervention, who in their judgement, would be "able to profit from the intervention" and they have been willing to exclude some of the lowest-achievers from selection' (Clay, 2005a, p. 22). Consequently, results for RR outcomes are likely to be inflated in terms of a central goal of RR: meeting the needs of the lowest achievers. Findings on the effect of placement in RR therefore, do not necessarily apply to the lowest achieving students, but rather some (unknown) number of students whose reading difficulties are milder.

Limitations of Reading Recovery in meeting the needs of struggling readers

Students who are successfully discontinued from RR tend to form the focus of published studies on the programme (D'Agostino et al., 2017). The characteristics of unrecovered students are largely unknown. Given this dearth of information, we turned to New Zealand annual monitoring data for information about students who exit the programme successfully, and those who are unrecovered and 'referred on' for further literacy learning assistance. We summarise data presented in the most recently available report, which is for the year 2015 (New Zealand Ministry of Education, 2016), to highlight the extent to which students benefit from RR, especially those students who most need early literacy supports. The lowest achieving students who were not placed in the RR programme cannot be included in these data, of course, because they are officially unknown.

The annual monitoring report includes two sources of information provided by schools: a school report at the end of each year, and reports on individual students (New Zealand Ministry of Education, 2016). The school report includes the number of students in RR and the number teachers and delivery hours allocated for RR for the year. Student reports present the background characteristics of students, time spent in RR, and entry and exit scores on three assessments: instructional text level, the Burt Word Reading Test—New Zealand Revision (Gilmore et al., 1981), and the Writing Vocabulary Task (Clay, 2002).

New Zealand uses levelled texts for use in junior primary school literacy programmes. A student's text level is defined as the level of books that she/he is able to read with 90–94% word recognition accuracy. This estimate is based on the results of a 'running record' (a form of miscue analysis) of the student's reading behaviour while reading a selected text. The Burt Word Reading Test measures context-free word recognition; students are asked to read words of increasing difficulty. This test correlates highly with word recognition in connected text ($r = .94$) and with reading comprehension ($r = .85$) (Blaiklock, 1997). In the Writing Vocabulary Task, students are allowed up to 10 minutes to write all the words they know on a blank sheet of paper. If necessary, a standard set of prompts is given.

In 2015, 60.3% of state schools with 6-year-old students offered the RR programme. These schools catered for 71.6% of the total New Zealand 6-year-old school population. Of the 6-year-old students in schools offering RR, 17.5% were placed in the programme in 2015. Notably, 24% of the students in RR schools were involved in RR at some stage during the year (which included RR students carried over from 2014). These percentages have been fairly stable for at least 13 years.

Socio-economic background factors are associated with both access to RR and success within the programme. New Zealand schools are rated on a decile system from 1 (low) to 10 (high) according to the socio-economic neighbourhood in which schools are located. The 2015 monitoring report showed that RR was more frequently offered in high-decile (8–10) schools (65% of high-decile schools) than in low-decile (1–3) schools (53% of low-decile schools). The percentages of students placed in RR varies in relation to school decile rankings. In high-decile schools, 10% of all 6-year-old children entered RR compared with 16% of all students enrolled in low-decile schools. Exacerbating this disparity, an average of 48 RR hours were allocated to each student in decile 1 schools, compared with 44 hours for each student in decile 10 schools. This difference resulted in approximately eight additional RR sessions for RR children in decile 1 schools. These data show that the schools that were more likely to have students requiring remedial assistance for reading (typically low-decile schools) were less likely to offer RR. When these schools did offer RR, students placed in the programme were typically there for a longer period of time and received more sessions.

In 2015, just over 36% of the total 6-year-old student population were Māori and Pasifika ('Pasifika' refers to people of Pacific Island Polynesian heritage born in the Pacific Islands or in New Zealand). These students made up nearly 47% of the students involved in RR. They are consistent with continuing trends over more than a decade. In the schools that offered RR, nearly 32% of Māori and Pasifika students were involved in the programme. These percentages are significantly higher than the 15.5% for New Zealand (NZ) European/Pākehā students (Pākehā is the commonly used Māori term for people of British or European (Caucasian) heritage). The higher RR participation rate for Māori and Pasifika students indicates that these students were already trailing their Pākehā peers in reading after only 1 year of formal schooling (in New Zealand children start school on or close to their fifth birthday).

Outcomes for RR in 2015 reveal that 78% of students were successfully discontinued, 14% were unrecovered and referred on for more specialist assistance, 6% left the school part way through the programme, and 3% did not continue in the programme

even though they were considered to be responding satisfactorily (New Zealand Ministry of Education, 2016, Table 10). Over the past 13 years (at least), 11–14% of RR children were unrecovered and referred on for specialist help (New Zealand Ministry of Education, 2016, p. 16). Fewer Māori and Pasifika students were successfully discontinued from RR than NZ European/Pākehā students (71%, 76% and 82% for Māori, Pasifika and NZ European/Pākehā students respectively). This pattern has been observed in RR outcome data since 2001. Over half (54%) of the students who were unrecovered and referred on were Māori and Pasifika.

We observed a similar pattern of outcomes for students attending low-decile schools. In decile 1–3 schools offering RR, 72% of students were successfully discontinued, compared with 83% of students in decile 8–10 schools. Concomitantly, 16% of students in decile 1–3 RR schools were unrecovered and referred on compared with 11% of students in decile 8–10 RR schools. A stark disparity exists between the unrecovered rate for students in decile 1 schools (18%) and those who were unrecovered in decile 10 schools (10%).

In summary, these data indicate that Māori and Pasifika students, and students from low-decile schools (often the same students) were less likely to have been recovered and more likely to have been referred on for specialist assistance. This pattern has been stable since at least the early 2000s, suggesting that the programme has not 'tailored' RR lessons adequately to meet the needs of Māori and Pasifika students and students from low-decile schools.

The reasons why such difficulties tend to occur with students from low socio-economic backgrounds is due to limited 'literate cultural capital' (Prochnow et al., 2015a). This term refers to literacy-related knowledge and abilities at school entry that result from activities in the home and pre-school environment that help with early literacy development (Tunmer & Nicholson, 2011). The knowledge, skills and strategies that students bring with them on entry to school include oral language ability, familiarity with 'book' or 'decontextualised' language and basic understanding of printed language, knowledge of letter names and sounds, ability to produce pre-conventional or invented spellings, phonological awareness and syntactic awareness (Prochnow et al., 2015a). Beginning readers who start school with limited amounts of literate cultural capital are more at risk for developing early reading difficulties. These students often have not had the pre-school language play activities and early literacy experiences that contribute to the development of crucial literacy-related skills. Constructivist approaches to literacy instruction, including RR, overlook such key differences among early readers and assume that all students have the necessary (middle class) language skills to benefit from the strongly text-oriented approach to literacy instruction.

The New Zealand Ministry of Education (2016) report also presents data on students' literacy learning gains from RR participation. Relative frequency distributions are shown for each of the three assessments (instructional text level, Burt Word Reading and Writing Vocabulary) on entry to and exit from RR, for discontinued and referred on students. The results show three especially significant patterns. First, referred on students had much lower mean (or modal) scores on the assessment measures on entry to RR than discontinued students. The modal entry score for instructional text level was 5 for discontinued students compared with 2 for referred

on students. For Burt Word Reading, discontinued students obtained a mean entry score of 13 compared with the referred on mean score of 6. For Writing Vocabulary, discontinued students had a mean entry score of 22 compared with 10 for the referred on students. Not surprisingly, this discrepancy was also true for exit scores: 18 versus 12 for instructional text level; 29 versus 17 for Burt Word Reading; and 60 versus 32 for Writing Vocabulary. Figures 1–3 illustrate the clear differences between Discontinued and Referred On students on the three measures. These mean entry and exit scores have remained virtually unchanged over the last few years. Students who entered RR with relatively low scores on the assessment measures were much more likely to be unrecovered after RR than students with higher entry scores.

These findings are consistent with Reynolds and Wheldall's (2007) conclusion that RR generally does not support the literacy learning needs of students who struggle the most with learning to read. Considering that unrecovered students are more likely to be Māori or Pasifika, and/or from low-income backgrounds, the results help to explain the failure of RR to significantly influence the persistently large achievement gap between good and poor readers in New Zealand. This gap includes a long tail of poor readers and has been a constant feature of literacy learning outcomes in New Zealand since the 1991 international study of literacy achievement that was undertaken by the International Association for the Evaluation of Education Achievement (IEA). The IEA study found that New Zealand had the largest spread of scores among the participating countries (Elley, 1992). The relatively large range of scores is a resilient characteristic of literacy learning outcomes in New Zealand. Four Progress in International Reading Literacy Study reports (PIRLS) since 2001 (Prochnow et al., 2015b; New Zealand Ministry of Education, 2017) show the consistency of the relatively large spread of scores. The RR programme was supposed to significantly reduce this gap, a claim that continues to be made on the New Zealand RR website: 'Reading Recovery is an effective early literacy intervention designed to significantly reduce the number of children with literacy difficulties in schools' (New Zealand Reading Recovery, 2018a, p. 1). The available evidence does not provide support for this claim.
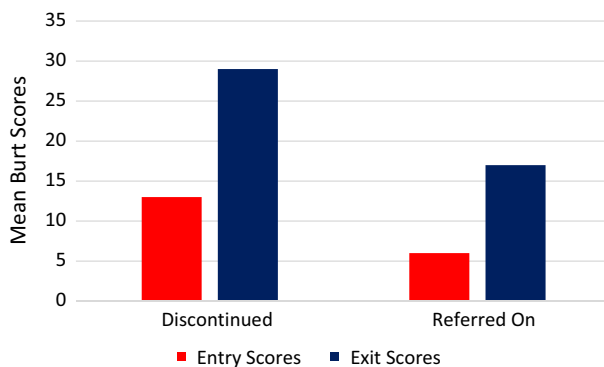

Figure 1.   Comparison of mean entry and exit Burt Word Test scores for discontinued and referred on students (derived from data presented in New Zealand Ministry of Education, 2016, pp. 26 and 27).
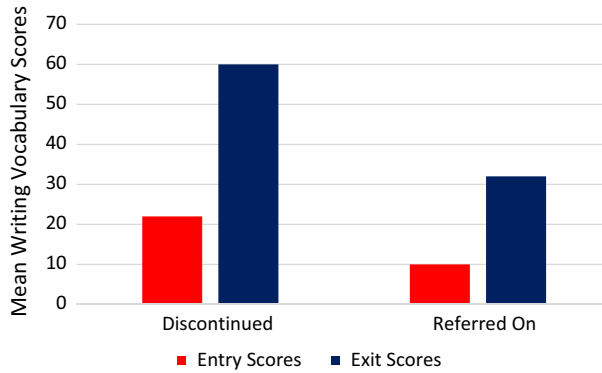
Figure 2.   Comparison of mean entry and exit writing vocabulary scores for discontinued and referred on students (derived from data presented in New Zealand Ministry of Education, 2016, pp. 29 and 30).
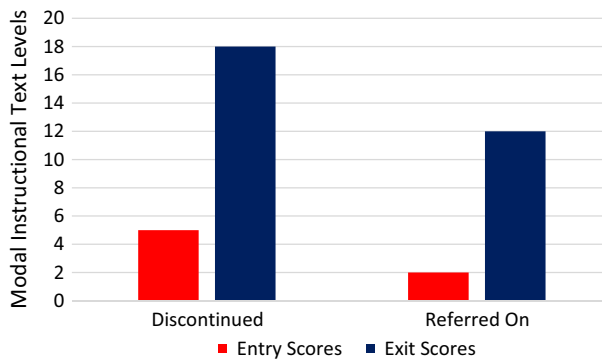

Figure 3.   Comparison of mean entry and exit instructional text levels for discontinued and referred on students (derived from data presented in New Zealand Ministry of Education, 2016, pp. 24 and 25).

The second noteworthy feature of RR students' gains data is the high variability in the RR entry and exit assessment scores. The entry and exit scores of Discontinued students overlapped so much for the Burt Word Test and the Writing Vocabulary Task that some students had entry scores that were greater than the mean of the exit scores. Figures 4 and 5 clearly illustrate the marked overlaps in the distributional pattern of scores for the Discontinued students. In addition, the figures illustrate the differences in the pattern of scores between the Discontinued and Referred On students.

These patterns are very unusual and reflect the fact that in New Zealand RR is offered to the lower (if not always the lowest) 15–20 percent of students within their same-age cohort in an individual school, regardless of the school's decile rating and associated levels of early literacy achievement. Because literacy achievement differences between students attending high- and low-decile schools are much larger in New Zealand than in most other countries (Mullis et al., 2007; Tunmer et al., 2008; Mullis et al., 2012; Prochnow et al., 2015b; Mullis et al., 2017), RR students in low-decile schools are much more likely to enter (and exit) the programme with lower scores on the assessment measures than RR students in high-decile schools. This finding accounts for the high level of variability in the RR assessment measures.
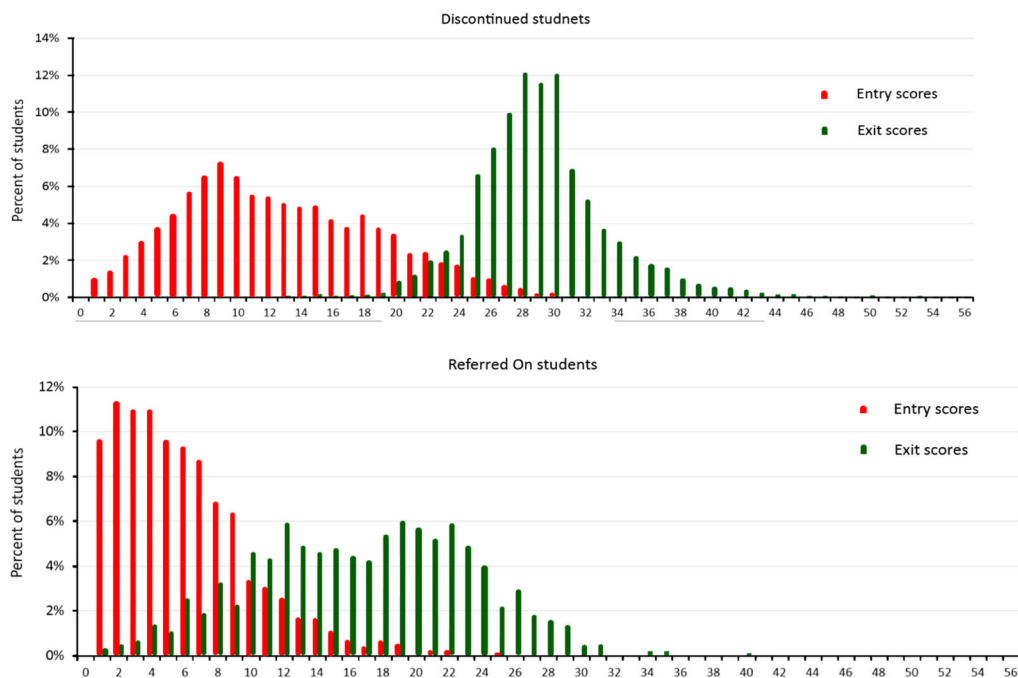
Figure 4. Comparative Burt Word Test distributional patterns of entry and exit scores for discontinued and referred on students (adapted from figures presented in New Zealand Ministry of Education, 2016, pp. 26 and 27).

The high level of variability has been a feature of RR assessments since at least the early 2000s, and probably since the start of the programme, because of the within-school selection process for placement in the programme. This selection model probably contributed to the relatively high false positive rate of approximately 30% of the students selected for participation in RR, as reported by Center et al. (1995). In other words, many RR students would have caught up with their peers without receiving RR instruction. As noted earlier, there appears to be a tendency to accept students into the RR programme who have more mild reading problems. Such problems can often be overcome spontaneously through regular classroom instruction and without the need for intervention.

The distributional pattern for instructional text level is the third significant feature of the data for entry and exit scores. The pattern of instructional text-level scores on exit from RR for Discontinued students was almost completely non-overlapping. The distribution of these scores was positively skewed and mainly truncated at text level 16, which is the recommended minimal level for discontinuation from RR (in 2015 approximately 1% of discontinued students had a text level of 15; New Zealand Ministry of Education, 2016, p. 24). Figure 6 illustrates the distributional pattern of text levels, comparing Discontinued students with Referred On students.

This truncated distributional pattern is in line with other New Zealand studies reporting marked discrepancies between RR teachers' and classroom teachers' assessments of text level at discontinuation (Glynn et al., 1989; Chapman et al., 2001). Problems associated with the use of RR teachers' assessments as an indicator of
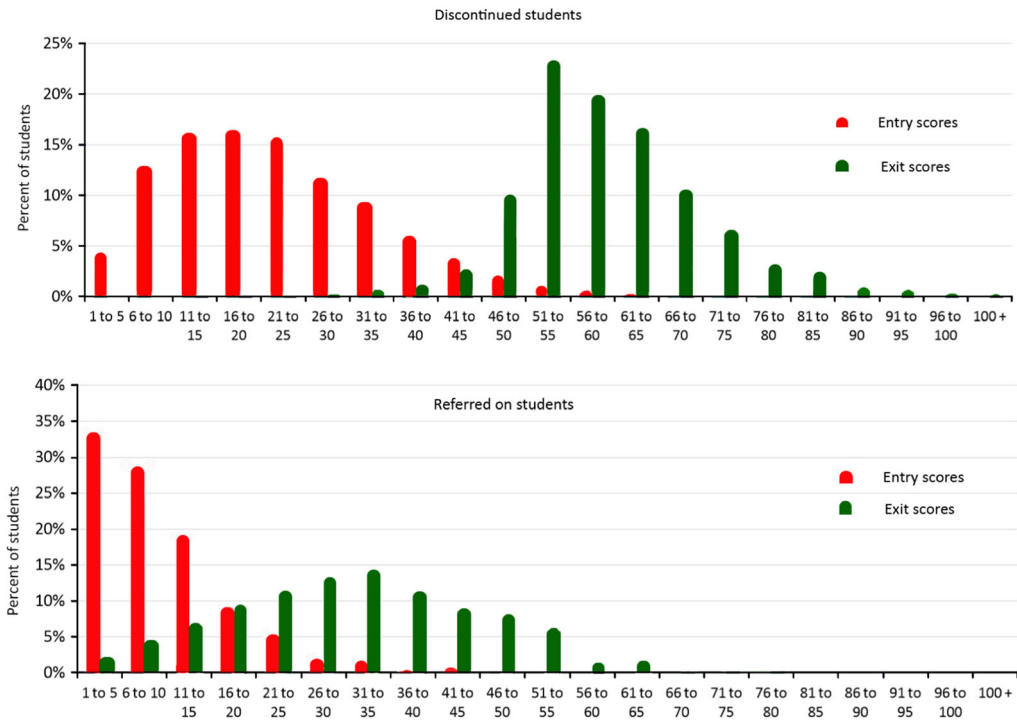
Figure 5.   Comparative Writing Vocabulary Test distributional patterns of entry and exit scores for discontinued and referred on students (adapted from figures presented in New Zealand Ministry of Education, 2016, pp. 29 and 30).

students' progress are seldom addressed. Assessments conducted by teachers who have a vested interest in a programme or teaching approach can lead to bias—intentional or unintentional (Reed et al., 2014). Classroom teachers' assessments of text level tend to be more accurate and more in line with independent measures of reading performance (such as the Burt Word Reading Test, which correlates highly with word recognition accuracy in connected text) than running records assessments by RR teachers (Glynn et al., 1989; Chapman et al., 2001). Moreover, text level, as determined by running records, is an unreliable form of assessment that produces inflated estimates of reading achievement, especially when undertaken by RR teachers (Leu, 1982; Allington, 1984; Hiebert, 1994; Elbaum et al., 2000; Chapman et al., 2001; Tunmer & Chapman, 2003; Blaiklock, 2004; Fawson et al., 2006).

In summary, data from the 2015 National Monitoring Report show that the participation rate in RR of Māori and Pasifika students is proportionately higher than for European/Pākehā students. Yet, Māori and Pasifika students along with students from low-income backgrounds are discontinued less often and more frequently unrecovered and referred on for additional specialist help. Low scores on the RR entry assessments appear to be predictive of whether or not the RR programme will be beneficial.

Perhaps the most significant feature of the National Monitoring Data since at least the early 2000s is the consistent percentages of students selected for the RR programme, their assessment scores, and the rates of referral for additional assistance. As
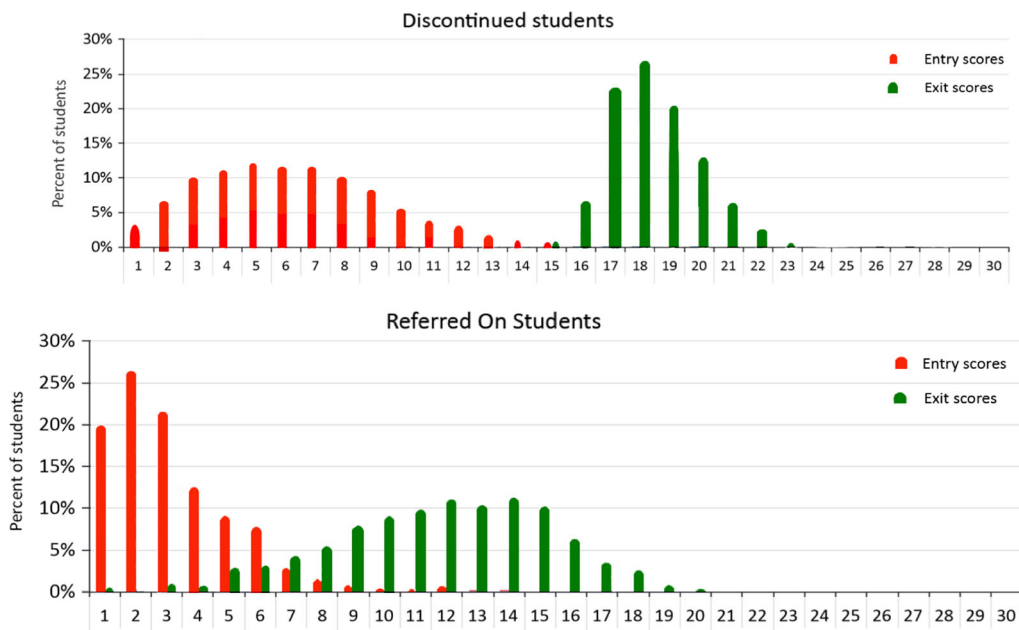
Figure 6.   Comparative Instructional Text distributional patterns of entry and exit levels for discontinued and referred on students (adapted from figures presented in New Zealand Ministry of Education, 2016, pp. 24 and 25).

noted elsewhere in relation to a 2012 RR National Monitoring report (Chapman et al., 2015), these stable patterns continue to include the contrasting percentages of Māori and Pasifika students in RR compared to NZ European/Pākehā students; disparate percentages of students in RR from high- and low-decile schools; consistent percentages of students who are discontinued or unrecovered; almost identical entry and exit mean scores for discontinued and also for unrecovered students; and consistent percentages of discontinuation and referral for additional support for Māori and Pasifika students compared to NZ European/Pākehā students. It is astonishing that these regular patterns have not been identified and addressed in order to increase the effectiveness of RR for all students, but especially for Māori and Pasifika students and those from low-decile schools.

This observation highlights the questionable claim by Clay that RR instruction is tailored to the individual needs of each student, with close attention being given to shifts in the student's responding and progress (Clay, 1998). There is no evidence in the patterns observed in these monitoring data that the programme is tailored to the needs of individual students, or that close attention is given to individual responding and progress. Had such individualised instruction truly been the case, the persistent patterns observed in RR outcomes would have been disrupted many years ago.

Further evidence of the failure of RR to meet the needs of those students who struggle the most with learning to read can be seen in the PIRLS 2011 data for students who had previously been in the RR programme. The PIRLS is a 5-year cycle of reading assessments that focuses on the reading achievement and literacy experiences of 9-year-old students throughout the world. In the 2011 survey, a specific New

Zealand question was on students' participation in a remedial programme (Question 8B). This question was included in the home-based 'Learning to Read Survey' presented to participating students' parents or caregivers. (The question was not included in the most recent [2016] PIRLS survey.) Slightly more than 60% of parents or caregivers (*n* = 3400) completed this home survey (Chamberlain, 2014, personal communication). Parents or caregivers were asked if their child had been in a remedial reading programme at any time from when they started school. If the response to this question was 'Yes', parents/caregivers were then asked to show from a list of three options the type of remedial programme their child had received. The RR programme was included in this list along with another (unspecified) school-based programme, or a programme that operated out of school hours (Chamberlain, 2014, personal communication).

Around 15% of parents/caregivers indicated that their children had received remedial reading assistance. RR was the programme most reported to having been received (*n* = 600; 69% of respondents to this question). This large percentage is not surprising because RR was and continues to be by far the most visible and widely available programme in New Zealand.

New Zealand students (*n* = 5600) obtained a PIRLS mean reading score of 531.02 (*SD* = 88.27). We analysed data to compare the mean score for students who had received RR with students who had not participated in any remedial programme. The result was a very large 75 point difference: RR *M* = 493.10, *SD* = 75.80; Non-Remedial *M* = 568.05, *SD* = 79.96. This difference between the two groups is equal to a negative effect size of around –0.96, meaning that around 83% of the RR group obtained PIRLS reading scores that were lower than the average student who was not placed in a remedial reading programme. It is useful to note that the mean score for students whose parents/caregivers did not complete the home survey was 502.55 (*SD* = 84.08), which is 9.5 points higher than for students who received RR. This means that non-responding parents/caregivers were not necessarily those who had students who required placement in a remedial reading programme because of low levels of literacy achievement.

In line with the 2015 National Monitoring Report data presented earlier, the PIRLS data showed significant differences as a function of school decile level. Former RR students who attended decile 1 or 2 schools obtained a mean reading score of 451.76 (*SD* = 75.70), whereas the mean reading score of those former RR students who attended decile 9 and 10 schools was 515.13 (*SD* = 75.80). Major differences in reading scores were also found as a function of ethnicity. The mean score for former NZ European/Pākehā RR students was 502.64 (*SD* = 80.59), compared to Māori students' mean score of 469.12 (*SD* = 72.55), and the Pasifika students' mean of 463.44 (*SD* = 71.20). In short, students who had received RR and who were Māori or Pasifika, or who were in decile 1 to 2 schools, had a mean PIRLS reading score that was approximately 100 points below that of students who did not receive any remedial assistance. This difference is equal to a negative effect size of over –1.0, meaning that at least 84% of former RR students who were Māori or Pasifika performed on PIRLS reading assessments at levels that were lower than the average for students who had not been placed in a remedial reading programme.

The PIRLS data relating to students' placement in the RR programme did not reveal which students were discontinued from RR and which were referred on for further assistance. The very consistent data presented in the RR National Monitoring Reports since the early 2000s show that around 80% of students are successfully discontinued from RR each year. This percentage indicates that a very large number of the former RR students in the 2011 PIRLS would have been discontinued from the programme. It is evident from the 2011 PIRLS data that these students, who would have received RR 3 years before the 2011 survey, obtained much lower reading scores than their cohort peers who were not placed in a remedial reading programme.

The significance of these PIRLS results lies in the role that RR was supposed to play as a prevention strategy against later literacy difficulties (New Zealand Reading Recovery, 2018a). The clear evidence from data in this PIRLS survey is that RR did not contribute to the prevention of later literacy difficulties. Given the stability of results from previous PIRLS surveys in 2001 and 2006 compared with the 2011 survey, it is highly probable that had the question been asked of parents/caregivers about their child's placement in a remedial reading programme, and particularly RR, the results would have been similar. Further, in the most recent (2016) PIRLS survey, results for New Zealand children showed a significant decline (Mullis et al., 2017). Former RR students in that recent survey would likely have performed no better than the former RR students in the 2011 survey, suggesting that RR continues to be unsuccessful in reducing ongoing literacy learning difficulties.

RR has also failed in an important goal of reducing the literacy learning gap in New Zealand, which is one of the largest gaps of all countries that participate in the PIRLS surveys (Prochnow et al., 2015b). Reducing this large gap, and in particular, reducing the disparity in reading achievement scores between Māori/Pasifika students and NZ European/Pākehā students, has been a major goal of successive governments in New Zealand since at least the late 1990s. RR was seen as a key element in this goal, as noted by New Zealand Reading Recovery: 'Reading Recovery is supported and significantly funded by the New Zealand Ministry of Education as part of the Literacy and Numeracy Strategy' (2018b, p. 1). We document the failure of this literacy strategy elsewhere (Prochnow et al., 2015b). Unfortunately, RR has not contributed to more successful literacy learning outcomes for students over a long period of time.

Considered together, the National Monitoring Report data and the 2011 PIRLS results for 9-year-old students who had received RR in Year 2 show that RR has not achieved its primary goals in New Zealand. Many struggling readers who receive the RR programme do not have their literacy learning needs met. The programme is especially ineffective for significant numbers of Māori and Pasifika students, and for students from schools in low socio-economic neighbourhoods. The PIRLS data also belie the claim by RR New Zealand on their website that the programme acts as an 'insurance against low literacy levels' (New Zealand Reading Recovery, 2018a, p. 1).

How to improve the effectiveness of Reading Recovery

The RR programme has been effective for some students. It has also been effective in terms of developing a framework for widespread implementation, control, and RR teacher training. Despite that, evidence from numerous sources shows that RR has

not addressed the needs of students who do not benefit from placement in RR. The reasons why the programme fails to meet the needs of students who are 'hardest to teach' include programme leaders overlooking the characteristics of students who enter and do not succeed in the programme—consistent since at least the early 2000s–as well as serious shortcomings in several aspects of RR. These shortcomings include the assessment battery that omits measures of phonological processing skills, the specific instructional strategies emphasised in the programme (e.g. the multiple cues approach to word identification), the type of programme delivery (one-to-one versus instruction in pairs or small groups), and the relationship between classroom literacy instruction and the RR programme. Major changes in all of these areas, along with changes in the theoretical framework of RR, would likely result in the programme being more effective for more students, both in terms of outcomes and cost (Tunmer & Chapman, 2003, 2004; Church, 2005; Reynolds & Wheldall, 2007).

Taking the issue of programme delivery, for example, Iversen et al. (2005) found that an early intervention programme based on RR could be delivered effectively to pairs of struggling readers. Although RR instruction in pairs required longer lessons (43 minutes versus 33 minutes), there were no major differences on any measures at discontinuation between students taught in pairs compared with those taught in one-to-one lessons. Thus, by increasing instructional time by around a quarter, RR teachers could double the number of students served without compromising the outcomes. Additional benefits were obtained in this study by supplementing the standard RR instructional approach with the teaching of explicit, out-of-context word analysis skills.

Schwartz et al. (2012) also found RR teacher–student ratios of greater than one-to-one did not sacrifice student learning outcomes. Results based on the Ohio Word Test (Pinnell et al., 1987) revealed that teacher–student ratios of 1:2 and 1:3 produced outcomes that did not differ from the 1:1 ratio. Results for the 1:2 ratio also did not differ from the 1:1 ratio for the Concepts About Print measure (Clay, 2002), Writing Vocabulary (Clay, 2002), the Slosson Oral Reading Test—Revised (Nicholson & Slosson, 2002), and the Hearing and Recording Sounds in Words measure (Clay, 2002). Both the Iversen et al. (2005) and Schwartz et al. (2012) studies indicate therefore, that RR teachers working with pairs of students were just as effective as working one-to-one, with considerable savings in time.

These findings are consistent with numerous studies relating to reading as well as to other subjects. Vaughn et al. (2010) found that one-to-three reading intervention was just as effective as one-to-one intervention. In the area of mathematics, Gervasoni (2005) found small group remedial intervention more effective than individualised tuition. Hattie (2009) cautioned against uncritical acceptance of the view that individualised instruction is better than small group instruction. He noted that evidence supporting individualised instruction is 'not so supportive' (p. 198). He cited research stressing the importance of instruction being adapted to the needs of students and that such instruction need not necessarily occur with individuals. These comments would suggest that individualised instruction may be little more than one-to-one instruction, with little effort to truly individualise (or tailor) instruction to support the specific needs of each student. RR teachers are trained to believe that individualised instruction is essential because this allows for the intervention to be 'tailored' to the

needs of each student. As we demonstrate in this review, there is little evidence to show that RR instruction is tailored to the individual needs of students. Rather, RR instruction is provided to individuals in a one-to-one setting, and presented according to an established RR view about each reading lesson.

Regarding the question of the relationship between RR and classroom literacy instruction, the programme was originally developed to complement regular whole language literacy instruction in New Zealand. Nevertheless, Clay (1993a) claimed that RR was compatible with all types of classroom literacy programmes, but she provided no evidence to support this claim. Center et al. (2001) tested this belief by investigating whether the effectiveness of RR varied as a function of the regular classroom literacy programme. They compared the effects of RR for students in whole language classrooms and those in 'code-oriented' classrooms (i.e. literacy instruction that included explicit and systematic instruction in phonological awareness and alphabetic coding skills). At the end of the second year of schooling, students in code-oriented classrooms (regular and RR students combined) significantly outperformed students in whole language classrooms on assessments of phonological recoding, reading connected text and invented spelling, as well as on a standardised test of reading comprehension.

Center et al. (2001) discussed a particularly important finding. RR students in the code-oriented classrooms also significantly outperformed the RR students in whole language classrooms on all four literacy measures. There was an average reading age advantage for the RR students in code-oriented classrooms of 8 months on the standardised reading comprehension measure. In addition, more RR students from whole language classrooms (83%) than from code-oriented classrooms (50%) were unrecovered (which was defined as falling below the 30th percentile on at least three of the four literacy assessments). RR students from whole language classrooms spent significantly greater time in the programme than students from code-oriented classrooms (18 vs 14.5 weeks). Further, far more RR students from whole language classrooms were either withdrawn from the programme or referred on for additional assistance compared with those from code-oriented classrooms (25% vs 5%). In general, however, the RR students in both types of classrooms were unable to reach the average level of their peers on any of the literacy measures. These findings refute Clay's (1993a) unsupported claim that the regular classroom context does not influence the literacy performance of RR students.

Although the nature of regular classroom instruction impacts the effectiveness of RR, the most serious inadequacy of the programme is the differential effectiveness at the individual level. Our central argument is that the RR programme is helpful for some struggling readers, but not those struggling readers who need help the most. For those students who experience considerable difficulties with learning to read, more intensive and systematic instruction in phonemic awareness and phonemically based decoding skills is likely to be required than what is normally provided in RR lessons (Iversen et al., 2005; Tunmer & Greaney, 2008, 2010).

These findings are consistent with the results of a major meta-analysis of one-to-one tutoring programmes for struggling readers. Slavin et al. (2011) found larger effect sizes for younger students in reading programmes that had more emphasis on phonics than those programmes that had less emphasis on phonics, including RR.

They noted that outcomes for RR were unexpectedly lower given the claims made for the success of the programme. A striking result from the Slavin et al. meta-analysis was the finding that the overall effect size for 18 studies involving paraprofessional or volunteer tutors using structured and intensive programmes was almost identical to the effect size for RR studies (+0.24 vs + 0.23), despite the much more intensive and expensive training that RR teachers receive. The one-to-one tutoring programmes that had a strong "phonetic" emphasis resulted in a substantially larger mean effect size of +0.62. Slavin et al. concluded that emphasising phonics greatly improves student literacy learning outcomes.

This conclusion echoes that reached by Hattie (2009). Based on the extensive meta-analyses documented in his book *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*, Hattie reported a mean effect size of +0.06 for whole language approaches to teaching reading (of which RR is one), compared with a mean effective size of +0.60 for phonics approaches. He concluded that 'phonics instruction is powerful in the process of learning to read—both for reading skills and for reading comprehension' (Hattie, 2009, p. 134).

The findings from Slavin et al. (2011) and Hattie (2009) provide strong support for the argument that explicit training in phonological decoding skills should be incorporated into the RR programme. Such a step, coupled with adoption of a 1:2 RR teacher to student ratio, would significantly increase the effectiveness of the programme, especially for those students who are in the very early phases of the developmental continuum for reading acquisition.

Making changes to the RR programme

These types of considerations and the wealth of research in support of such instruction, led the Literacy Experts Group set up by the Ministry of Education to advise a New Zealand Literacy Taskforce in the late 1990s (New Zealand Ministry of Education, 1999a), to reach the following unanimously agreed upon recommendation: 'We recommend that Reading Recovery places greater emphasis on explicit instruction in phonological awareness and the use of spelling-to-sound patterns in identifying unfamiliar words in text' (New Zealand Ministry of Education, 1999a, p. 6). The Literacy Taskforce (New Zealand Ministry of Education, 1999b) rejected this recommendation. Instead, the Taskforce recommended a review of the RR programme (p. 23).

Making significant changes to RR following a review, however, would have been difficult. Rivers (2001), who interviewed Clay, reported that:

> If any changes were made to Reading Recovery, they could be made to its administration only, or they could risk being in breach of the programme's trademark. Its developer, Marie Clay, said she held a trademark on the name Reading Recovery to protect the programme's integrity. (p. 1)

The RR programme continues to be protected. In New Zealand, RR is overseen by the Marie Clay Literacy Trust, which is responsible for the copyright of all RR materials and the RR trademark. Changes cannot be made to the materials or procedures without the trustees' approval (Marie Clay Literacy Trust, 2018). Although the Ministry of Education funds the programme in New Zealand, control over the

programme by the Marie Clay Literacy Trust makes it virtually impossible for changes based on contemporary scientific evidence to be made to the RR programme, or even to undertake independent studies identifying suitable modifications to the programme to improve student outcomes and cost effectiveness.

Despite the overwhelming research on the benefits of including systematic attention to phonological processing and word-level decoding skills in early intervention tutoring programmes for struggling readers, proponents of RR consider problems with the programme lie with implementation issues and not the instructional framework. In a report that was favourable to RR, McDowall et al. (2005) found that the programme was less helpful to Māori and Pasifika students than to other students, which is a finding that is consistent with material we have presented in this paper. McDowall et al. reported that RR teachers and RR tutors interviewed in the study attributed problems associated with the relatively poor benefits of RR for Māori and Pasifika students to implementation, resourcing, family/cultural factors, and inappropriate textual materials. Problems with the nature of the programme itself were not considered. McDowall et al. overlooked the fundamental drawback of RR, which is that it is based on the multiple cues/searchlights theory of reading—a model of reading that was rejected by the scientific community in the early 1980s (e.g. Stanovich, 1980). It is bewildering that those responsible for oversight of the RR programme continue to resist changes to the theoretical framework, materials, and teaching approach. This resistance is difficult to comprehend given the wealth of scientific research, monitoring data, and follow-up studies, all of which identify significant shortcomings with this programme.

Conclusion

Slavin (2016) asserted that RR is 'One of the very, very few unquestioned success stories of evidence-based reform' (p. 61). The view that RR has attained 'unquestioned success' is contrary to an abundance of research on RR showing otherwise. Our review of literature and analyses of RR data from annual monitoring reports indicate that RR is of some benefit for some students, at least in the short term, but of limited benefit to those students who most need help during the early stages of literacy acquisition.

Students who derive limited or no benefit from placement in the RR programme have less well developed foundation skills required for successful literacy acquisition. These students tend to be in the pre-alphabetic or partial-alphabetic phases identified by Ehri (2005). They typically have limited or severely limited phonemic awareness and phonemically based decoding skills on entry to RR (Chapman et al., 2001). Further, as shown in data from the New Zealand National Monitoring Report for 2015, students who are unrecovered from RR have significantly lower entry scores on measures of instructional text level, context-free word identification, and writing vocabulary (New Zealand Ministry of Education, 2016). This pattern of low RR entry scores for unrecovered students, along with low exit scores, has been consistent and well known since at least the early 2000s, and probably since the start of the programme.

Unrecovered RR students are likely to attend schools located in low SES neighbourhoods, and in New Zealand are often Māori or Pasifika. Such students typically

have limited literate cultural capital, and frequently do not possess the language skills that the RR instructional model appears to take for granted (Prochnow et al., 2015a). Because RR does not provide adequate, systematic and explicit instruction in phonological awareness and the use of letter–sound relations, the programme fails to meet the needs of these students, and does not 'tailor' instruction in accordance with their entry skills and developmental levels.

Little, if anything, appears to have been done by those responsible for RR to address the needs of children at risk for remaining unrecovered following placement in the programme. Yet, there is ample evidence to indicate that students who struggle the most with learning to read benefit from systematic, explicit instruction in phonemic awareness and the use of letter–sound relations, along with strategy training in how and when to use this knowledge while reading text and to spell words while writing messages. Adopting these instructional approaches could enhance the RR programme and lead to greater benefits for larger numbers of students.

Concerns about the limitations of the RR programme led to a comment in a recent report of the New Zealand Education Review Office (ERO) that teachers should not rely on the RR programme to adequately address the literacy learning needs of students who struggle with learning to read (New Zealand Education Review Office, 2018). The ERO report noted that there is an assumption 'that all students that have not progressed initially when learning to read will participate in RR and will succeed' (p. 37). However, the report draws attention to the shortcomings of the RR programme for Māori and Pasifika students, and for students attending schools in low SES neighbourhoods, as well as to evidence that gains attained by some students in RR are often not sustained. Because of the level of investment in RR in New Zealand, the ERO report recommends that the issues need investigating. The concerns expressed in the ERO report about RR represent the first time since the inception of the programme in New Zealand that a state sector agency has questioned the long-held beliefs and assumptions about the efficacy of the programme.

Similar concerns about the efficacy of RR contributed to the decision by the New South Wales Department of Education to discontinue supporting the programme in that Australian state. According to Smith (2017), the RR programme in New South Wales costs A\$50 million to run, but the programme is ineffective, and improvements for some students are often short-lived.

The ongoing failure of RR to meet the needs of those children who most require literacy learning supports appears to stem from the failure of those responsible for the programme to accept changes to RR based on contemporary scientific research. If the RR programme cannot be changed to reflect contemporary scientific research on reading interventions, it should be dropped and replaced by a more contemporary, research-based reading intervention approach.

References

Allington, R. L. (1984) Oral reading, in: P. D. Pearson, R. Barr, M. L. Kamil & P. B. Mosenthal (Eds) *Handbook of reading research* (vol. 1) (New York: Longman), 829–864.

Belgrave, J. W. (2009) *The characteristics of children who are referred on from Reading Recovery*. Unpublished Master of Education thesis (Hamilton, New Zealand, University of Waikato).

Blaiklock, K. (2004) The importance of letter knowledge in the relationship between phonological awareness and reading, *Journal of Research in Reading*, 27(1), 36–57.

Blaiklock, K. E. (1997) Bring back the Burt: Some comments on the value of word recognition tests for the assessment of reading, *New Zealand Reading Forum*, 2, 13–16.

Center, Y., Freeman, L. & Robertson, G. (2001) The relative effect of a code-oriented and a meaning-oriented early literacy program on regular and low progress Australian students in year 1 classrooms which implement Reading Recovery, *International Journal of Disability, Development and Education*, 48, 207–232.

Center, Y., Wheldall, K., Freeman, L., Outhred, L. & McNaught, M. (1995) An evaluation of Reading Recovery, *Reading Research Quarterly*, 30(2), 240–263.

Chamberlain, M. (2014) Personal communication, 24 August, 2014.

Chapman, J. W., Greaney, K. T. & Tunmer, W. E. (2007) How well is Reading Recovery really working in New Zealand? *New Zealand Journal of Educational Studies*, 42(1–2), 17–29.

Chapman, J. W., Greaney, K. T. & Tunmer, W. E. (2015) Is Reading Recovery an effective early literacy intervention programme for children who most need literacy supports? in: W. E. Tunmer & J. W. Chapman (Eds), *Excellence and equity in literacy education: The case of New Zealand* (Basingstoke, UK, Palgrave Macmillan), 41–70.

Chapman, J. W. & Tunmer, W. E. (2011) Reading Recovery: Does it work? *Perspectives on Language and Literacy*, 37(4), 21–24.

Chapman, J. W. & Tunmer, W. E. (2016) Is Reading Recovery an effective intervention for students with reading difficulties? A critique of the i3 scale-up study, *Reading Psychology*, 37(7), 1025–1042. https://doi.org/10.1080/02702711.2016.1157538.

Chapman, J. W., Tunmer, W. E. & Prochnow, J. E. (2001) Does success in the Reading Recovery program depend on developing proficiency in phonological-processing skills? *Scientific Studies of Reading*, 5(2), 141–176.

Church, J. (2005, December) *Accelerating reading development in low achieving children: A review of research*, paper presented at the annual meeting of the New Zealand Association for Research in Education, Dunedin, New Zealand.

Clay, M. M. (1979) *The early detection of reading difficulties* (2nd edn) (Auckland, New Zealand, Heinemann).

Clay, M. M. (1980) Reading Recovery: A follow-up study, *New Zealand Journal of Educational Studies*, 15(2), 137–155.

Clay, M. M. (1985) *The early detection of reading difficulties* (3rd edn) (Auckland, New Zealand, Heinemann).

Clay, M. M. (1987) Learning to be learning disabled, *New Zealand Journal of Educational Studies*, 22(2), 155–173.

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

Clay, M. M. (1993a) *Reading Recovery: A guidebook for teachers in training* (Auckland, New Zealand, Heinemann).

Clay, M. M. (1993b) *An observation survey of early literacy development* (Auckland, New Zealand, Heinemann).

Clay, M. M. (1998) *An observation survey of early literacy achievement* (Auckland, New Zealand, Heinemann).

Clay, M. M. (2002) *An observation survey of early literacy achievement* (2nd edn) (Auckland, New Zealand, Heinemann).

Clay, M. M. (2005a) *Literacy lessons designed for individuals. Part One: Why? When? And How?* (Auckland, New Zealand, Heinemann).

Clay, M. M. (2005b) *Literacy lessons designed for individuals. Part Two: Teaching procedures* (Auckland, New Zealand, Heinemann).

Clay, M. M. (2007) International perspectives on the Reading Recovery program, *Journal of Reading Recovery*, Fall, 16–34.

Cook, P., Rodes, D. R. & Lisitz, K. L. (2017) The reading wars and Reading Recovery: What educators, families and taxpayers should know, *Learning Disabilities: A Multidisciplinary Journal*, 22 (2), 12–23. https://doi.org/10.18666/ldmj-2017-v22-i2-8391

D'Agostino, J. V. & Harmey, S. J. (2016) An international meta-analysis of Reading Recovery, *Journal of Education for Students Placed at Risk*, 21(1), 29–46.

D'Agostino, J. V., Lose, M. K. & Kelly, R. H. (2017) Examining the sustained effects of Reading Recovery, *Journal of Education for Students Placed at Risk*, 22(2), 116–127.

D'Agostino, J. V. & Murphy, J. A. (2004) A meta-analysis of Reading Recovery in United States schools, *Educational Evaluation and Policy Analysis*, 26(1), 23–28.

Darr, C., McDowall, S., Ferral, H., Twist, J. & Watson, V. (2008) *Teacher manual: Progressive achievement test: reading* (Wellington, New Zealand, New Zealand Council for Educational Research).

Ehri, L. C. (2005) Development of sight word reading: Phases and findings, in: M. J. Snowling & C. Hulme (Eds) *The science of reading: A handbook* (Oxford, UK, Blackwell), 135–154.

Elbaum, B., Vaughn, S., Hughes, M. & Moody, S. (2000) How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research, *Journal of Educational Psychology*, 92(4), 605–619.

Elley, W. B. (1992) *How in the world do students read?* (Hamburg, Germany, International Association for the Evaluation of Educational Achievement).

Fawson, P. C., Ludlow, B. C., Reutzel, D. R., Sudweeks, R. & Smith, J. A. (2006) Examining the reliability of running records: Attaining generalizable results, *The Journal of Educational Research*, 100, 113–126.

Gervasoni, A. (2005) The diverse learning needs of young children who were selected for an intervention program, in: H. L. Chick & J. L. Vincent (Eds) *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (vol. 3) (Melbourne, Australia, PME), 33–40. Retrieved from https://pdfs.semanticscholar.org/efab/913f09f27bf2ad1f0546ebd40a94c607a7a6.pdf

Gilmore, A., Croft, C. & Reid, N. (1981) *Burt word reading test: New Zealand revision* (Wellington, New Zealand, New Zealand Council for Educational Research).

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

Greaney, K. T. (2011) The multiple cues or 'searchlights' word reading theory: Implications for Reading Recovery, *Perspectives on Language and Literacy*, 19(2), 15–19.

Grossen, B., Coulter, G. & Ruggles, B. (1999) *Reading Recovery: An evaluation of benefits and costs*. Unpublished paper, University of Oregon. Retrieved from http://darkwing.uoregon.edu/~bgrossen/rr.htm

Hattie, J. (2009) *Visible learning: A synthesis of over 800 meta-analyses relating to achievement* (London, Routledge).

Hiebert, E. H. (1994) Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher*, 23(9), 15–25.

Hollands, F. M., Pan, Y., Shand, R., Cheng, H., Levin, H. M., Belfield, C. R., Kieffer, M., Bowden, A. B. & Hanisch-Cerda, B. (2013) *Improving early literacy: Cost-effectiveness analysis of effective reading programs* (Center for Benefit-Cost Studies of Education, Teachers College, Columbia University). Retrieved from: http://cbcse.org/publications/

Iversen, S. A. & Tunmer, W. E. (1993) Phonological processing skill and the Reading Recovery program, *Journal of Educational Psychology*, 85(1), 112–125.

Iversen, S., Tunmer, W. E. & Chapman, J. W. (2005) The effects of varying group size on the Reading Recovery approach to preventative early intervention, *Journal of Learning Disabilities*, 38(5), 456–472.

Jesson, R. & Limbrick, L. (2014) Can gains from early literacy interventions be sustained? The case of Reading Recovery, *Journal of Research in Reading*, 37(1), 102–117.

Leu, D. J. (1982) Oral reading analysis: A critical review of research and application, *Reading Research Quarterly*, 17(3), 420–437.

Marie Clay Literacy Trust. (2018) http://irrto.org/marie-clay-literacy-trust/

May, H., Gray, A., Sirinides, P., Goldsworthy, H., Armijo, M., Sam, C., Gillespie, J. N. & Tognatta, N. (2015) Year one results from the multisite randomized evaluation of the i3 scale-up of reading recovery, *American Educational Research Journal*, 52(3), 547–581.

McDowall, S., Boyd, S. & Hodgen, E. (2005) *Reading Recovery in New Zealand: Uptake, implementation, and outcomes, especially in relation to Māori and Pasifika students* (Wellington, New Zealand, New Zealand Council for Educational Research).

McLachlain, K. (2014) Further thoughts on Reading Recovery, *Learning Disabilities Australia Bulletin*, 46, 39.

McNaughton, S. (2007) Dame Professor Marie M. Clay: Scientist, professor, colleague. *Journal of Reading Recovery*, 7(1), 104–106.

Mullis, I. V. S., Martin, M. O., Foy, P. & Drucker, K. T. (2012) *PIRLS 2011 international results in reading* (Chestnut Hill, MA, TIMSS & PIRLS International Study Center, Boston College). Retrieved from http://timss.bc.edu/pirls2011/international-results-pirls.html

Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (Eds) (2017) *PIRLS 2016 international results in reading* (Chestnut Hill, MA, Boston College, TIMSS & PIRLS International Study Center). Retrieved from https://timssandpirls.bc.ed/pirls2016/international-results/

Mullis, I. V. S., Martin, M. O., Kennedy, A. M. & Foy, P. (2007) *PIRLS 2006 International Report* (Chestnut Hill, MA, TIMSS & PIRLS International Study Center, Boston College).

New Zealand Education Review Office. (2018) *Evaluation at a glance: A decade of assessment in New Zealand primary schools—practice and trends* (Wellington, New Zealand, Author). Retrieved from: http://www.ero.govt.nz/assets/Uploads/Evaluation-at-a-glance-assessment-practice-and-trends-2020.pdf

New Zealand Ministry of Education. (1999a) *Literacy Experts Group: Report to the Secretary for Education* (Wellington, New Zealand, Author).

New Zealand Ministry of Education (1999b) *Report of the Literacy Taskforce* (Wellington, New Zealand, Author).

New Zealand Ministry of Education. (2016) *Annual monitoring of Reading Recovery: The data for 2015* (Wellington, New Zealand, Author). Retrieved from: http://www.educationcounts.govt.nz/__data/assets/pdf_file/0004/180769/2015-Annual-Monitoring-of-Reading-Recovery.pdf

New Zealand Ministry of Education. (2017) *PIRLS 2016: New Zealand's achievement* (Wellington, New Zealand, Author). Retrieved from: https://www.educationcounts.govt.nz/__data/assets/pdf_file/0007/183697/PIRLS-2016-NZs-Achievement.pdf

New Zealand Reading Recovery. (2018a) *Reading Recovery*. Retrieved from https://www.readingrecovery.ac.nz/reading_recovery/

New Zealand Reading Recovery. (2018b) *Index*. Retrieved from https://www.readingrecovery.ac.nz/index.php

Nicholas, K. & Parkhill, F. (2013) Is Reading Recovery sustainable two to four years after discontinuation? *Education 3-13: International Journal of Primary, Elementary and Early Years Education*, 42(6), 663–672.

Nicholson, C. L. & Slosson, R. L. (2002) *Slosson Oral Reading Test - Revised (SORT-R)* (East Aurora, NY, Slosson Educational Publications).

Nicholson, T. (1989) A comment on Reading Recovery, *New Zealand Journal of Educational Studies*, 24, 95–97.

Pinnell, G., Lyons, C., Young, P. & DeFord, D. (1987) *Reading Recovery in Ohio, Volume VI*. Technical Report (Columbus, OH, The Ohio State University).

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. & Seltzer, M. (1994) Comparing instructional models for the literacy education of high at-risk first graders, *Reading Research Quarterly*, 29(1), 8–39.

Pressley, M. (2006) *Reading instruction that works: The case for balanced teaching* (New York, NY, Guilford Press).

Prochnow, J. E., Tunmer, W. E. & Arrow, A. W. (2015a) Literate cultural capital and Matthew effects in reading achievement, in: W. E. Tunmer & J. W. Chapman (Eds) *Excellence and equity in literacy education: The case of New Zealand* (Basingstoke, UK, Palgrave Macmillan), 145–167.

Prochnow, J. E., Tunmer, W. E. & Greaney, K. T. (2015b) Findings from New Zealand's participation in the Progress in International Reading Literacy Study in 2001, 2006, and 2011, in: W. E. Tunmer & J. W. Chapman (Eds) *Excellence and equity in literacy education: The case of New Zealand* (Basingstoke, UK, Palgrave Macmillan), 23–40.

Reading Recovery Council of North America. (2009). *Standards and guidelines of Reading Recovery in the United States* (5th ed.). Columbus, OH: Author. Retrieved from http://www.readingrecovery.org/pdf/implementation/Standards_Guidelines-09_Full_Version.pdf

Reading Recovery Council of North America. (2018) https://readingrecovery.org/reading-recovery/teaching-children/basic-facts/

Reed, D. K., Cummings, K. D., Schaper, A. & Biancarosa, G. (2014) Assessment fidelity in reading intervention research: A synthesis of the literature, *Review of Educational Research*, 84(2), 275–321.

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.

Reynolds, M., Wheldall, K. & Madelaine, A. (2009) The devil is in the detail regarding the efficacy of Reading Recovery: A rejoinder to Schwartz, Hobsbaum, Briggs, and Scull, *International Journal of Disability, Development and Education*, 56(1), 17–35.

Rivers, J. (2001, February 16) Reading Recovery review, *New Zealand Education Review*, 1.

Schwartz, R. M. (2005) Literacy learning of at-risk first grade students in the Reading Recovery early intervention, *Journal of Educational Psychology*, 97(2), 257–267.

Schwartz, R. M., Schmitt, M. C. & Lose, M. K. (2012) Effects of teacher–student ratio in response to intervention approaches, *The Elementary School Journal*, 112(4), 547–567.

Serry, T., Rose, M. & Liamputtong, P. (2014) Reading Recovery teachers discuss Reading Recovery: A qualitative investigation, *Australian Journal of Learning Difficulties*, 19(1), 61–73.

Shanahan, T. (1987) The shared knowledge of reading and writing, *Reading Psychology: An International Quarterly*, 8(2), 93–102. https://doi.org/10.1080/0270271870080205.

Shanahan, T. (2017) Reading research: The importance of replicability, *The Reading Teacher*, 70 (4), 507–510.

Shanahan, T. & Barr, R. (1995) Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk readers, *Reading Research Quarterly*, 30(4), 958–996.

Slavin, R. E. (2016) Getting to scale: Evidence, professionalism, and community, *Journal of Education for Students Placed at Risk*, 21, 60–63.

Slavin, R. E., Lake, C., Davis, S. & Madden, N. A. (2011) Effective programs for struggling readers: A best evidence synthesis, *Educational Research Review*, 6(1), 1–6.

Smith, A. (2017, 31 December) New literacy teachers recruited as NSW government axes Reading Recovery, *The Sydney Morning Herald*. Retrieved from https://www.smh.com.au/education/new-literacy-teachers-recruited-as-nsw-government-axes-reading-recovery-20171222-h09c5f.html

Smith, J. W. A. & Elley, W. B. (1994) *Learning to read in New Zealand* (Auckland, New Zealand, Longman Paul).

Snow, C. E. & Juel, C. (2005) Teaching children to read: What do we know about how to do it? in: M. J. Snowling & C. Hulme (Eds) *The science of reading: A handbook* (Oxford, UK, Blackwell), 501–520.

Soler, J. & Openshaw, R. (2006) *Literacy crises and reading policies* (London, Routledge).

Stanovich, K. (1980) Toward an interactive-compensatory model of individual differences in the development of reading fluency, *Reading Research Quarterly*, 16(1), 32–71.

Tunmer, W. E. & Chapman, J. W. (2003) The Reading Recovery approach to preventive early intervention: As good as it gets? *Reading Psychology*, 24(3–4), 337–360.

Tunmer, W. E. & Chapman, J. W. (2004) Reading Recovery: Distinguishing myth from reality, in: R. M. Joshi (Ed.) *Dyslexia: Myths, misconceptions, and some practical applications* (Baltimore, MD, International Dyslexia Association), 99–114.

Tunmer, W. E. & Greaney, K. T. (2008) Reading intervention research: An integrative framework, in: G. Reid, A. Fawcett, F. Manis & L. Siegel (Eds) *The Sage handbook of dyslexia* (London, Sage), 241–267.

Tunmer, W. E. & Greaney, K. T. (2010) Defining dyslexia, *Journal of Learning Disabilities*, 43(3), 229–243.

Tunmer, W. E., Greaney, K. T. & Prochnow, J. E. (2015) Pedagogical constructivism in New Zealand literacy education: A flawed approach to teaching reading, in: W. E. Tunmer & J. W. Chapman (Eds) *Excellence and equity in literacy education: The case of New Zealand* (Basingstoke, UK, Palgrave Macmillan), 121–144.

Tunmer, W. E. & Nicholson, T. (2011) The development and teaching of word recognition skill, in: M. L. Kamil, P. D. Pearson, E. B. Moje & P. Afflerbach (Eds) *Handbook of reading research* (vol. 4) (New York, Routledge), 405–431.

Tunmer, W. E., Nicholson, T., Greaney, K. T., Prochnow, J. E., Chapman, J. W. & Arrow, A. W. (2008) PIRLS before swine: A critique of New Zealand's national literacy strategy, *New Zealand Journal of Educational Studies*, 43(2), 105–119.

University College London. (2018) http://www.ucl.ac.uk/international-literacy/reading-recovery

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2007, August) *Beginning reading*. Retrieved from https://education.ucf.edu/mirc/research/ies%20-%20beginning%20reading.pdf

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, July) *Beginning reading intervention report: Reading Recovery*. Retrieved from http://whatworks.ed.gov

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, October) *WWC review of the report: Evaluation of the i3 scale-up of Reading Recovery year one report, 2011–2012*. Retrieved from http://whatworks.gov

Vaughn, S., Denton, C. A. & Fletcher, J. M. (2010) Why intensive interventions are necessary for students with severe reading difficulties, *Psychology in the Schools*, 47(5), 432–444.

Vellutino, F. R. & Fletcher, J. M. (2005) Developmental dyslexia, in: M. J. Snowling & C. Hulme (Eds) *The science of reading: A handbook* (Oxford, UK, Blackwell), 362–378.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J. & Scanlon, D. M. (2004) Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45(1), 2–40.

Wasik, B. A. & Slavin, R. E. (1993) Preventing early reading failure with one-to-one tutoring: A review of five programs, *Reading Research Quarterly*, 28(2), 179–200.

Wilkinson, I. A. G., Freebody, P. & Elkins, J. (2000) Reading research in Australia and Aotearoa/New Zealand, in: M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds) *Handbook of reading research* (vol. 3) (Mahwah, NJ, Erlbaum), 3–16.

Wilson, K. G. & Daviss, B. (1994) *Redesigning education* (New York, Teachers College Press).