WHY THE NEW ZEALAND NATIONAL LITERACY STRATEGY HAS FAILED AND WHAT CAN BE DONE ABOUT IT

Evidence from the Progress in International Reading Literacy Study (PIRLS) 2011 and Reading Recovery Monitoring Reports

William E. Tunmer
James W. Chapman
Keith T. Greaney
Jane E. Prochnow
Alison W. Arrow

Massey University Institute of Education

July 2013

CONTENTS

Page

SUMMARY

In this paper arguments and evidence are presented showing that New Zealand's national literacy strategy has failed, why it has failed, and what can be done to overcome the problem.

The paper begins with a summary of the key reports, reviews, and government policy initiatives that have occurred over the past 15 years to indicate the levels of concern that educators and policy makers have expressed regarding the persistently large inequities in literacy achievement outcomes.

New Zealand's relatively "long tail" of literacy underachievement was a major concern for educators and policy makers that grew during the 1990s. In response to these growing concerns, the Government established a Literacy Taskforce to provide recommendations aimed at raising the literacy achievement of all students but with particular attention given to "closing the gap between the lowest and highest students" (Ministry of Education, 1999b, p. 7).

The recommendations of the Taskforce constituted the *national literacy strategy* for reducing the large disparity in reading achievement outcomes between good and poor readers. The Ministry of Education (MoE) was given the responsibility of implementing the recommendations of the Taskforce.

More recently, in December 2011, the MoE's *Briefing to the Incoming Minister* (Ministry of Education, 2011) stated that:

> … the gap between our high performing and low performing students remains one of the widest in the Organization of Economic Cooperation and Development (OECD). These low performing students are likely to be Māori or Pasifika and/or from low socio-economic communities. Disparities in education appear early and persist throughout learning. (p. 8)

Based on these findings the *Briefing* concluded that, **"The greatest challenge facing the schooling sector is producing equitable outcomes for students"** (p. 23).

The key question arising from the summary of policy initiatives and government reports on literacy achievement over the past 15 years is this:

> **Why have concerns expressed by policy makers, teachers' unions, and legislative bodies regarding New Zealand's relatively large literacy achievement gap continued for such a long period of time? The answer is that New Zealand's national literacy strategy is simply not working.**

The remainder of the paper is divided into three sections. The first section presents findings showing that New Zealand's national literacy strategy has failed. The second section presents arguments and evidence regarding the major factors responsible for the persistence of New Zealand's wide gap in literacy achievement, and for why the gap has not diminished over the past 15 years despite major efforts by the MoE to address the problem. The third section reviews research on the most effective strategies for reducing the literacy achievement gap.

Evidence that New Zealand's National Literacy Strategy has Failed

Evidence that New Zealand's national literacy strategy has failed is demonstrated in two recently released reports: The Progress in International Reading Literacy Study (PIRLS) 2011 report (Mullis, Martin, Foy, & Drucker, 2012), and the latest annual monitoring report of Reading Recovery (RR) data in New Zealand, which includes trend data for the past 10 years (Lee, 2011).

PIRLS 2011

The PIRLS focuses on the achievement and literacy learning experiences of children from countries throughout the world in grades equivalent to Year 5 in New Zealand. It is a 5-year cycle of assessments that was first administered in 2001, then in 2005/2006, and again in 2010/2011. It includes a state-of-the-art test of reading comprehension that was designed to assess two aspects of reading literacy: purposes of reading and processes of comprehension.

The PIRLS 2011 results show that attempts by the Ministry of Education to reduce the large disparity between good and poor readers have failed, despite a decade of policies and resources aimed at closing the gap. Virtually no changes in educational outcomes have occurred.

- The average reading achievement score for New Zealand in the PIRLS 2011 study was not significantly different from either the PIRLS 2001 or 2006 studies (see Table 1).
- The number of countries that significantly outperformed New Zealand exceeded the number of countries that New Zealand significantly outperformed (as occurred in the PIRLS 2006 study).
- Of the six English-speaking comparison counties, all but one significantly outperformed New Zealand.
- Trend data revealed that, although there were more increases than decreases in mean reading achievement scores across countries from 2001 to 2011, New Zealand showed no significant increases in reading performance.
- The standard deviation (variation in scores) and range (between the 5th and 95th percentiles) for New Zealand's reading scores were almost unchanged from the PIRLS 2001 and 2006 studies and exceeded the values of most other countries, including those of the six English-speaking comparison countries.
- The large differences in reading achievement scores between Pākehā/European and Māori/Pasifika students have also not changed over the past decade (see Table 2).
- There were no significant changes from the PIRLS 2001 or 2006 results in either the relatively high percentage of New Zealand students who performed at the advanced international benchmark or the relatively high percentage of students who failed to reach the low international benchmark, despite a general improvement across other countries in the percentages of students reaching international benchmarks from 2001 to 2011.

Reading Recovery 2011

Regarding RR, if the programme had been successful in achieving its goal of substantially reducing the number of children who develop ongoing reading difficulties, then the large gap in reading performance consistently observed between good and poor readers since the 1991

international study of literacy achievement by the International Association for the Evaluation of Educational Achievement (Elley, 1992) should have steadily decreased after RR was introduced throughout the country in the mid-1980s. This has not been the case.

Data from RR annual monitoring reports and other sources indicate that RR has had little or no impact on reducing New Zealand's relatively large literacy achievement gap.

The reason is that the RR programme is of limited benefit to those students who need help the most, especially Māori/Pasifika students and students from low-income backgrounds. The data show that RR is less accessible to Māori and Pasifika students and students from low-income backgrounds, probably because of RR's relative ineffectiveness for these students. Also, a significant number of the lowest performing 6-year-olds are excluded from RR because they are considered unlikely to benefit from the programme, or they are withdrawn early because RR teachers could not bring them up to the expected rates of progress.

Māori and Pasifika students and students from low-income backgrounds are less likely to be successfully discontinued from RR and more likely to be referred on for specialist help, even though they generally receive extra lessons and spend more time in the programme than Pākehā students.

Students who enter RR with relatively high scores on the assessment measures of RR are much more likely to benefit from RR than students with relatively low scores. This finding adds to the evidence that RR generally does not work well for students who are most at risk for failing to learn to read.

Finally, research indicates that positive maintenance effects for the majority of successfully discontinued RR students are modest or non-existent. For these reasons RR has had little or no impact on reducing New Zealand's relatively large literacy achievement gap.

Why has New Zealand's National Literacy Strategy Failed?

Three factors have contributed to the failure of New Zealand's national literacy strategy:

- a constructivist orientation toward literacy education,
- the failure to respond adequately to differences in literate cultural capital at school entry, and
- restrictive policies regarding the first year of literacy teaching.

Pedagogical constructivism in literacy education

New Zealand has followed a predominantly constructivist approach to literacy education for the past 25 years. In this approach literacy learning is largely seen as the *by-product* of active mental engagement. There is little or no explicit, systematic teaching of phonemic awareness (the ability to reflect on and manipulate the phonemic segments of spoken words) and alphabetic coding skills (the ability to translate letters and letter patterns into phonological forms). Yet, both phonemic awareness and alphabetic coding skills are essential for learning to read successfully.

Underpinning the constructivist approach to literacy teaching is the "multiple cues" theory of reading (sometimes called the "searchlights" model). According to this view, skilled reading is a process in which minimal word-level information is used to confirm predictions about the upcoming words of text based on multiple sources of information (Clay, 1991). Learning to read is seen largely as a process in which children learn to use multiple cues in identifying words in text. Text-based cues (i.e., picture cues, sentence context cues, preceding passage context, prior knowledge activated by the text) are used by students to predict the text yet to be encountered. Letter-sound information is generally used only to confirm word predictions or guesses and for self-correction (Clay, 1998).

The scientific community has firmly rejected the constructivist/multiple cues model of reading (Pressley, 2006). The major shortcoming of the multiple cues approach is that it stresses the importance of using information from many sources in identifying unfamiliar words in text without recognizing that skills and strategies involving phonological information are of primary importance in beginning literacy development.

Research indicates that for progress to occur in learning to read, **the beginning reader must acquire the ability to translate letters and letter patterns into phonological forms** (Ehri, 2005; Snow & Juel, 2005; Tunmer & Nicolson, 2011). To discover mappings between spelling patterns and sound patterns, children must also be able to segment spoken words into subcomponents. A large body of research shows that explicit, systematic attention to alphabetic coding skills in early reading instruction is more effective than non-systematic or no phonics instruction that characterize constructivist approaches (Brady, 2011; Hattie, 2009; Snow & Juel, 2005; Tunmer & Arrow, 2013).

RR is also based on the multiple cues theory of reading. In RR lessons particular emphasis is placed on reading strategies that involve children developing the flexible use of multiple cues to detect and correct errors while reading text (Clay, 2005a, b). Although there are serious shortcomings and much needed improvements in several aspects of RR, the most serious shortcoming concerns the differential effectiveness of the program.

The programme is beneficial for some struggling readers but not others, especially those struggling readers who need help the most. Research indicates that for these children, more intensive and systematic instruction in phonemic awareness and phonemically-based decoding skills is needed than what is normally provided in RR lessons (Chapman, Tunmer, & Prochnow, 2001; Church, 2005; Iversen, Tunmer, & Chapman, 2005).

Based on such findings, the Literacy Experts Group (Ministry of Education, 1999a) that advised the Literacy Taskforce (Ministry of Education, 1999b) included in its report 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 (p. 6).

This recommendation was rejected by the Literacy Taskforce.

Literate cultural capital, Matthew effects in reading achievement, and PIRLS 2011

For the past 20 years New Zealand has consistently shown comparatively high levels of variability in the test scores from international surveys of reading achievement. We have argued that New Zealand's relatively wide spread of scores is largely the result of Matthew (rich-get-richer and poor-get-poorer) effects (Stanovich, 1986). These Matthew effects are triggered by a predominantly constructivist, multiple cues approach to reading instruction and intervention that fails to respond adequately to differences in literate cultural capital at school entry (Tunmer, Chapman, & Prochnow, 2003). Literate cultural capital is a generic term referring to literacy-related knowledge and abilities at school entry that come from activities in the home environment that support early literacy development (Tunmer & Nicholson, 2011).

Children who do not have sufficient levels of essential reading-related skills when they start formal reading instruction (and who are not provided with supplementary instruction to develop these competencies, especially phonological awareness) are forced to rely increasingly on ineffective word identification strategies. Continued reliance on ineffective strategies, such as picture cues, partial visual cues, and contextual guessing, inevitably leads to literacy learning difficulties (Pressley, 2006). Because of their ineffective word identification skills, these children not only receive less practice in reading but eventually come across reading materials that are too difficult for them. **This often results in avoidance of reading, poor attention in class, low expectations of success, and withdrawal from literacy learning tasks (i.e., negative Matthew effects).**

Reliance on these poor word identification strategies is exacerbated by the multiple cues approach to teaching reading because emphasis is placed on encouraging beginning readers to use text-based cues (i.e., semantic-contextual and syntactic-contextual cues). Not enough attention is given to the development of phonemically-based word-level skills and strategies that are essential for success in learning to read. Lack of attention to these word-level skills is a significant disadvantage for those children who are less able to discover letter-sound patterns as a by-product of more general reading (Torgesen, 2004).

There is further evidence that New Zealand's large literacy achievement gap can be explained in terms of literate cultural capital and a constructivist orientation toward literacy education. Research indicates that:

- children enter school with large individual differences in the skills and competencies (i.e., literate cultural capital) important in learning to read;
- children from low-income and/or culturally diverse backgrounds on average begin school with considerably lower levels of literate cultural capital than middle-class children;
- children who possess higher levels of literate cultural capital at the beginning of school generally profit more from literacy instruction, learn to read sooner, and read better than those who do not;
- according to data from the PIRLS 2001, 2006, and 2011 studies, differences in literate cultural capital at the beginning of school are associated with larger differences in future reading achievement in New Zealand than in most other countries.

Given these findings, the challenge for policy makers in the MoE is to develop an approach to literacy education in which the new entrant with limited literate cultural

capital has approximately the same chance of success in learning to read as the new entrant with an abundance of literate cultural capital.

Restrictive policies regarding the first year of literacy teaching

New Zealand educators and policy makers are very resistant to providing beginning readers with assessment and explicit instruction in skills that are essential for reading development (e.g., phonological awareness, alphabetic coding skills), especially during the first year of schooling when instruction in these skills would be most effective. This resistance stems from three sources.

Reading Recovery. The first formal assessment of literacy skills in New Zealand occurs at the end of the child's first year of schooling with the use of the Observation Survey developed by Clay (1998). Clay (2005a) argued that this in-depth assessment should not occur until the end of the child's first year of formal instruction because "the child should be given sufficient time to adjust to the school situation and a variety of opportunities to pay attention to literacy activities" (p. 12). *However*, research has shown that a more effective strategy for improving reading among struggling readers is to intervene at an *earlier* point (Lonigan & Phillips, 2012). Wagner (2008) argued against a "wait-to-fail" approach to reading intervention. Instead, new entrants should receive an initial evaluation consisting of measures of emergent literacy skills that are known to be important in early literacy development (e.g., phonological awareness, print awareness). Supplementary instruction in these skills would then be given to those children who needed it.

Constrained Skills Theory (CST). Constrained skills include phonological awareness, alphabetic coding skills, and reading fluency (i.e., automaticity in word recognition). Skilled readers ultimately master these skills completely and in a relatively short period of time. In contrast, unconstrained skills develop more slowly and are never completely mastered, as they continue to develop over the course of a lifetime. Unconstrained skills include vocabulary and comprehension skills.

Although not explicitly stated, the MoE appears to be using the distinction between constrained and unconstrained skills to avoid including in the national literacy standards an assessment of the core skills that underlie the development of literacy (Greaney & Tunmer, 2010). Instead, the literacy standards document focuses largely on describing sets of comprehension-enhancing strategies that students should be able to demonstrate at each year level. **Teachers are not required or encouraged to undertake systematic assessments of constrained skills at any point during the primary school years.** The Observation Survey (Clay, 1998) administered to children after they have completed a year of formal schooling does not include measures of phonological awareness, alphabetic coding skills, or reading fluency.

The MoE has highlighted a quote from Paris (2005) relating to constrained and unconstrained skills in providing the theoretical basis for national literacy standards. But the MoE needs to consider more carefully what Paris (2005) actually said about constrained skills. He stated that beginning readers "need to be instructed on those skills *early and persistently* by teachers and parents. Constrained skills *must be mastered*" (p. 199, emphasis added). Paris further said that "constrained skills need to be mastered because they are necessary but not sufficient for other reading skills. They enable automatic decoding, deployment of attention, and

application of comprehension strategies so they *set the stage for reading development*" (p. 200, emphasis added).

These comments by Paris (2005) strongly suggest that the MoE needs to place much greater emphasis on the assessment and teaching of constrained skills (i.e., phonological awareness, alphabetic coding skills, automaticity in word recognition) during the first year of formal schooling.

Culturally responsive instruction. Another source of the strong resistance to early skills-based teaching is the assumption that this approach to reducing the large inequities in achievement outcomes in literacy education is based on "deficit theory" that "pathologizes" children from low-income, culturally diverse backgrounds, and is antithetical to culturally responsive instruction (Berryman & Bishop, 2011; Fayden, 2005; Harris, 2009). Harris, for example, stated that words such as *gap*, *underachievement*, *disparity*, and *at-risk* "signal perceived deficiencies" (p. 12).

The views expressed by Harris (2009) and others appear to question the overwhelming research evidence that learning to read is a developmental process that takes place over time, involves qualitatively different (but perhaps overlapping) phases, and may break down at different points due to the failure to acquire the core skills that underlie the development of literacy (Ehri, 2005; Pressley, 2006; Snow & Juel, 2005; Tunmer & Nicolson, 2011).

The literacy learning needs of children necessarily vary because they differ in: (i) the amount of reading-related knowledge, skills, and experiences (i.e., literate cultural capital) they bring to the classroom; (ii) the explicitness and intensity of instruction they require to learn skills and strategies for identifying words and comprehending text, and; (iii) their location along the developmental progression from pre-reader to skilled reader. **It is better to talk in terms of the *specific learning needs* of students, not deficiencies.**

Teachers can be effective in reducing the gap in literacy achievement if they address the specific needs of children struggling to learn to read. Teachers can do this *regardless* of the cultural group or social class to which children belong. There is no evidence that explicit instruction in essential reading-related skills (i.e., constrained skills) cannot be done in a culturally sensitive and responsive manner. Teachers should certainly adjust their teaching to support students' identities, while at the same time holding the same expectations and standards of achievement for students of diverse backgrounds as for those from the dominant culture.

What can be done to Overcome the Failure of New Zealand's National Literacy Strategy?

Little or no progress has been made in reducing the literacy achievement gap because the constructivist/multiple cues model of reading adopted by the MoE as the theoretical basis for its approach to literacy teaching and intervention is fundamentally flawed.

The strategies for reducing the large inequities in achievement outcomes in New Zealand literacy education are to:

1. Make fundamental changes to regular classroom literacy instruction.

2. Replace RR with an alternative intervention programme that is specifically designed to target those struggling readers who need help the most.

Classroom literacy instruction

To reduce the unacceptably large gap in literacy achievement in New Zealand, classroom literacy instruction will need to change to reduce the influence of differences in literate cultural capital at school entry on future reading achievement. The most effective strategy for reducing the literacy achievement gap is to use differentiated instruction from the outset of formal schooling that takes into account interactions between school entry reading-related skills (high vs. low literate cultural capital) and method of teaching reading (constructivist vs. explicit approaches).

For some beginning readers, the processes of acquiring literacy skills are highly *learner dependent*. These children seem to grasp the idea of what is required to discover orthographic patterns after having had only a small amount of phonologically-based skills and strategies explicitly taught to them.

In contrast, for other children the learning processes are more *environment dependent.* These children require a fairly structured and teacher-supported introduction to reading (Snow & Juel, 2005).

Research indicates that these differences in literacy learning processes depend largely on the amount of literate cultural capital that children possess at school entry. *Learner-dependent* children have higher levels of essential reading-related knowledge, skills, and experiences, and *environment-dependent* children have more limited amounts (Juel & Minden-Cupp, 2000).

A predominantly constructivist, book experience approach to reading instruction with a major emphasis on reading books, writing text, and some incidental teaching of word analysis skills during reading and writing activities is therefore likely to be more suitable for *learner-dependent* children than heavy code-emphasis approaches.

However, *environment-dependent* children will almost certainly benefit more from beginning reading instruction that includes explicit, systematic teaching of phonological awareness and alphabetic coding skills outside the context of reading text *in combination with* plenty of opportunities to practice and receive feedback on using these skills during text reading.

In support of these claims, Connor, Morrison, and Katch (2004) found that children who began first grade with below-average reading-related skills made larger reading gains in classrooms that provided greater amounts of teacher-managed, code-focused instruction throughout the year than in classrooms that provided greater amounts of child-managed, meaning-focused instruction. In contrast, for children with high levels of reading-related skills at school entry, greater growth in reading was achieved in classrooms that provided lesser amounts of teacher-managed, code-focused instruction and greater amounts of child-managed, meaning focused instruction.

Of particular importance was the finding that when student characteristics were appropriately matched with instructional approach, the improvement in end of year reading scores resulting

from good fitting instructional patterns varied greatly between children with high and low levels of literate cultural capital at the beginning of first grade.

For high literate cultural capital children, better fitting instructional patterns (i.e., child-managed, meaning focused instruction) resulted in about half a grade equivalent gain in end of year reading scores over less effective instructional patterns.

However, for low literate cultural capital children, better fitting instructional patterns (i.e., teacher managed, code-focused instruction) resulted in a difference of more than *two full grade equivalents* in end of year reading scores compared with poorly fitting instructional patterns.

These findings have particular significance for literacy education in New Zealand. The relatively high level of disparity in reading achievement outcomes among New Zealand readers can be explained in terms of a largely unidimensional, constructivist approach to literacy teaching that produces poorly fitting instructional patterns for students with limited amounts of literate cultural capital at the beginning of school.

The results of the Connor et al. (2004) study indicate that instructional strategies that may be effective with some students may be less effective for other students with different skills. In support of this claim, Connor et al. (2009) reported that children in first-grade classrooms that individualized reading instruction by taking into account child-by-instruction interactions made greater gains in reading achievement than children in classrooms that did not have individualized programs.

Reading Recovery

The RR programme is currently overseen by the Marie Clay Literacy Trust, which is responsible for the copyright of all RR materials and the RR trademark. No changes in the materials or procedures of RR can therefore be made without approval of the trustees. This makes it virtually impossible for school systems or countries (including New Zealand) to make changes to the RR programme based on recent research *or* to conduct independent studies investigating ways of modifying the programme to improve outcomes and/or cost effectiveness.

In a study of RR, McDowall, Boyd, Hogden, and van Vliet (2005) found that RR was less beneficial for Māori and Pasifika students than for other students. Problems associated with the benefits of RR for Māori and Pasifika were generally attributed to implementation, resourcing, family/cultural factors, and inappropriate textual materials *but not to the programme itself*.

McDowall et al. overlooked the fundamental problem with RR, which is that it is based on the multiple cues theory of reading, a model of reading that was rejected by the scientific community over three decades ago (e.g., Stanovich, 1980). As Church (2005) noted, RR "was designed in the 1970s prior to most of the modern research into how children learn to read. Not surprisingly, therefore, it lacks a number of elements which have been found by research to be essential in teaching low achieving children how to read" (p. 13).

As part of the effort to overcome the failure of New Zealand's national literacy strategy, RR needs to be replaced with an intervention programme that is based on

contemporary theory and research on reading intervention and targets children who are most at risk of failing to learn to read.

Conclusions

The constructivist/multiple cues model of reading adopted by the Ministry of Education as the theoretical basis for its approach to literacy teaching and intervention is fundamentally flawed. **As a consequence, little or no progress has been made in reducing New Zealand's unacceptably large inequities in literacy achievement outcomes.**

This claim is supported by the PIRLS 2011 results which show that virtually no changes in educational outcomes have occurred over the past decade. The mean reading achievement score for New Zealand in the PIRLS 2011 study was not significantly different from the PIRLS 2001 or 2006 studies, and the wide distribution of New Zealand's reading scores was almost unchanged from the PIRLS 2001 and 2006 studies. This wide distribution, manifest in the "long tail" of low scores, continues to be greater than most other countries.

Analyses of Reading Recovery data from annual monitoring reports indicated that RR has had little or no impact on reducing New Zealand's relatively large literacy achievement gap because the programme is of limited benefit to those children who need help the most, especially Māori/Pasifika students and students from low-income backgrounds. In addition, research indicates that positive maintenance effects for the majority of those students who are successfully discontinued from the RR programme, are modest or non-existent.

The arguments and evidence presented in this study should not be taken as suggesting that New Zealand teachers are responsible for the failure of the national literacy strategy. Data from the PIRLS 2011 study indicate that New Zealand teachers are well above the international average in level of formal education and availability of school resources for teaching reading.

Rather, the failure is largely the result of the misguided policy decisions of the Ministry of Education over the past 15 years. Doing the same thing (in terms of literacy instruction) with more resources, will only lead to the same, unacceptable results.

INTRODUCTION

Is the New Zealand approach to literacy education as successful as often claimed? For example, Smith and Elley (1997), two leading New Zealand literacy educators, noted that "expert commentators from other countries have been fulsome in their praise of our reading programmes, our reading teachers, our reading materials and our Reading Recovery methods" (p. 110). They further stated that "our methods of teaching . . . are all spreading to other parts of the world" and that, "It is no wonder that New Zealand is held up as the country whose reading programmes are 'best in the world' (*Newsweek*, 1991)" (p. 110). More recently, in a report on the New Zealand education system from the Center on International Education Benchmarking (2012), reference was made to "New Zealand's world-class work in the field of reading instruction," work which has "paid off handsomely" (p. 5).

In this paper we question these widely held assumptions. We begin by presenting a brief summary of the key reports, reviews, and government policy initiatives that have occurred since the publication of Smith and Elley's (1997) influential book describing New Zealand's approach to teaching reading. We then present arguments and evidence indicating that New Zealand's national literacy strategy has failed, why it has failed, and what can be done to solve the problem.

LITERACY REPORTS, REVIEWS, AND GOVERNMENT POLICY INITIATIVES

For the past 15 years the New Zealand government has initiated major efforts to reduce the relatively large inequities in achievement outcomes in literacy education. One of the first studies to draw attention to the relatively high levels of disparity between good and poor readers in New Zealand schools was the international study of literacy achievement carried out in 1991 by the International Association for the Evaluation of Educational Achievement (IEA). The results showed that New Zealand had the largest spread of scores among the participating countries (Elley, 1992) and that the low performing readers were likely to be Māori (the indigenous people of New Zealand) and/or from low-income backgrounds (Wagemaker, 1993). Further research in New Zealand during the 1990s revealed disparities between children of different backgrounds in important literacy-related skills at school entry (Gilmore, 1998; Nicholson, 1997) and that differences in literacy achievement between Māori and New Zealand European (Pākehā) students steadily increased over the first years of schooling (Crooks & Caygill, 1999; Flockton & Crooks, 1997), throughout high school (Nicholson, 1995; Nicholson & Gallienne, 1995), and into adulthood (Ministry of Education, 1997a). Home language was not considered as a possible explanation of the lower mean literacy achievement scores of Māori students because a small number of Māori learn to speak Māori as a first language (Crooks & Caygill, 1999).

Given the important role that literacy skills have in determining children's educational and life chances, the growing body of evidence of New Zealand's relatively "long tail" of literacy underachievement became a major source of concern among educators and policymakers in the 1990s and was regarded by leading reading researchers as the "single biggest challenge confronting literacy education in New Zealand today" (Wilkinson, Freebody, & Elkins, 2000, p.8). In response to these growing concerns, a Literacy Taskforce was established by the Government to provide advice on achieving its goal that, "By 2005, every child turning nine will be able to read, write, and do maths for success" (Ministry of Education, 1999b, p. 4). To assist the Government in developing an effective national literacy strategy, the Taskforce, which comprised mostly practitioners, focused on recommendations aimed at raising the

literacy achievement of all students but with particular attention given to "closing the gap between the lowest and highest achievers" (p. 7).

In addition to the Taskforce, a Literacy Experts Group was convened "to provide the Taskforce with advice from a range of theoretical and academic perspectives on literacy learning" (Ministry of Education, 1999a, p.1). The Literacy Experts Group comprised literacy researchers from New Zealand tertiary institutions and the New Zealand Council for Educational Research. Because not all of the recommendations made by the Literacy Experts Group were adopted by the Literacy Taskforce, the Literacy Experts Group submitted its own report which included several recommendations not made by the Taskforce (Ministry of Education, 1999a). For example, the Literacy Experts Group recommended that "greater attention needs to be focused on the development of word-level skills and strategies in beginning reading instruction, including the development of phonological awareness" (p. 6).

In March 2000, the Education and Science Committee of the New Zealand Parliament initiated an inquiry into the teaching of reading in New Zealand to determine "how and why many children are failing to learn to read effectively" and "to provide recommendations to the Government on how the reading gap can be closed" (New Zealand House of Representatives, 2001, p. 5). Following the inquiry, the Committee made 51 recommendations which were largely rejected by the Government, especially those calling for significant changes in New Zealand's approach to literacy education, such as the recommendations "that the Ministry of Education provide advice and support to schools to incorporate successful phonics programmes into the classroom" (p. 17), "that all primary teacher-training providers incorporate the teaching of phonetic skills and word-level decoding into their programmes" (p. 27), and that "there be a greater emphasis on the benefits of phonics instruction in Literacy Leadership materials" (p. 28). Instead, the government decided to adopt the recommendations of the Literacy Taskforce (Ministry of Education, 1999b). For example, the Taskforce recommended that "a description of the knowledge, skills, and attitudes that nine-year-olds demonstrate when they are reading and writing for success . . . be developed and promulgated to teachers and parents" (p. 9), such as that "a nine-year-old reading for success can predict, check, confirm, and self-correct while they are reading" (p. 32). Other recommendations included drawing up and promulgating to schools a statement of best practice in literacy instruction, developing a video "that illustrates taking and analysing running records . . . and using this [*sic*] data to inform the teaching programme" (p. 17), developing a comprehensive professional development package "to assist teachers to implement best practice in their teaching of reading and writing" (p. 19), and developing a nationally co-ordinated system of reading interventions that reviews and builds "on the interventions that already exist, in particular, Reading Recovery and the Resource Teachers of Reading" (p. 23). The recommendations of the Taskforce constituted the *national literacy strategy* for reducing the large disparity in reading achievement outcomes between good and poor readers.

The Ministry of Education (MoE) was given the responsibility of implementing the recommendations of the Taskforce. As part of this effort, the MoE established the Literacy Reference Group in May, 2006, to provide informed advice and guidance on future directions of the New Zealand Literacy Strategy. The Group, which comprised mostly practitioners, met at least once each year until 2011 when the Group was disbanded. During this time the Group discussed the MoE's Literacy Strategy Progress Report (for internal discussion only), the draft Literacy Learning Progressions (Ministry of Education, 2010), the draft National Literacy Standards (Ministry of Education, 2009), strategies for assisting students with

literacy learning problems, and the findings of the Progress in International Reading Literacy Study (PIRLS) 2006 (Mullis, Martin, Kennedy, & Foy, 2007).

The results of the PIRLS 2006 study became a major source of concern to the MoE and the Literacy Reference Group because virtually no reduction in the relatively large disparity between good and poor readers had occurred since the PIRLS 2001 assessment (Mullis, Martin, Gonzalez, & Kennedy, 2003). One of the MoE's arguments for the continuation of the unacceptably large literacy achievement gap reported in the PIRLS 2001 study was that the Government's literacy strategy "was introduced to schools in 2000, a year before the PIRLS 2001 survey, and possibly too early to have impacted on the PIRLS 2001 results" (Ministry of Education, 2003b, p.8). However, the MoE could not use this argument to account for the disappointing results from the PIRLS 2006 study, as the Government's literacy strategy had been largely operational during the period following the PIRLS 2001 study (Tunmer, Nicholson, Greaney, Prochnow, Chapman, & Arrow, 2008).

The PIRLS 2006 results contributed to two further developments. In March 2006, the Education and Science Committee of Parliament initiated an inquiry into "making the schooling system work for every child" (New Zealand House of Representatives, 2008, p.37). In support of the decision to conduct the inquiry, the Committee cited a recent report by the Education Review Office stating that "New Zealand's best students perform with the best in other countries but there is a group at the bottom, perhaps as large as 20 percent, who are currently not succeeding in our education system" (p. 7). The 20 percent estimate represented an aggregation of data from various sources, including the PIRLS reports. From these reports and submissions received during 2006 and 2007, the Committee concluded that "evidence from national and international assessments and studies support the proposition that New Zealand has a disproportionate number of students who underachieve" (p. 7). Recommendations of the Committee included devoting more resources to the "provision of comprehensive professional development in assessment practice so that by 2010 all schools will have experienced appropriate training in the collection and use of data" (p. 3). Unlike the 2001 report of the Education and Science Committee, no recommendations were made regarding New Zealand's approach to teaching reading.

The second development prompted by the PIRLS 2006 results was a symposium organized by the New Zealand Educational Institute (NZEI; the primary school teachers' union) on May 13, 2008, "to have an open and frank discussion on why the results from PIRLS are as they are" (S. Aikin, personal communication, April, 2008). The symposium included educational researchers, professional development providers, and representatives of the teaching profession (Chamberlain, 2012). All participants in the symposium were invited to make recommendations to the NZEI for modifying literacy teaching practices. Some participants suggested making major changes (such as ensuring the development of phonemically-based word-level skills and strategies by all beginning readers, and replacing the Reading Recovery early intervention programme with an alternative programme that is based on contemporary theory and research on reading intervention) but these recommendations were subsequently rejected by the broader NZEI membership. Following the symposium the NZEI, with the support of the MoE, initiated and funded a number of focus group meetings with teachers throughout New Zealand in 2008 and 2009 to discuss the PIRLS findings "from a New Zealand perspective" and to make the findings "more accessible and relevant to teachers" (Chamberlain, 2012, p. 456).

In 2010 the MoE introduced national standards in reading and writing for Years 1 through 8 as another strategy for reducing the literacy achievement gap (Ministry of Education, 2009). For Years 1 through 3 the standards are based on the book levels of the Ready to Read series, the core instructional series of books for New Zealand students. For example, the reading standard after one year at school is that "students will read, respond to, and think critically about fiction and non-fiction texts at the Green level of Ready to Read" (p. 20). Students meeting this standard are expected to read seen texts at the Green level with at least 90% accuracy (the Green level corresponds to a reading age of approximately 6 years). For each of Years 1 through 8, the reading standards also include illustrated examples of reading behaviours that teachers would be expected to observe in students who are meeting the standard.

More recently, in December 2011, the MoE's *Briefing to the Incoming Minister* (Ministry of Education, 2011), which occurs when a new government is formed after a national election, stated that, although there have been some overall improvements in education (largely in participation and retention rates):

> … the gap between our high performing and low performing students remains one of the widest in the Organization of Economic Cooperation and Development (OECD). These low performing students are likely to be Māori or Pasifika and/or from low socio-economic communities. Disparities in education appear early and persist throughout learning. (p. 8)

The *Briefing* indicated that over the past decade there has been little improvement in early literacy/numeracy, especially for Māori and Pasifika children (*Pasifika* refers to Polynesian descendants from Pacific Islands, such as Samoa, Tonga, and the Cook Islands). Data presented in the *Briefing* show that 18% of Māori and 16% of Pasifika are not achieving basic literacy and numeracy skills by age 10, compared to only 4% of Non-Māori and Non-Pasifika children (p. 9). Based on these findings the *Briefing* concluded that, "The greatest challenge facing the schooling sector is producing equitable outcomes for students" (p. 23). Improving the quality of teaching, placing greater emphasis on the accountability framework for schools, and establishing charter schools were identified as strategies that would be pursued by the MoE to improve achievement outcomes.

The important question that arises from our brief summary of policy initiatives and government reports on literacy achievement in New Zealand over the past 15 years is this: Why have concerns expressed by policy makers, teachers' unions, and legislative bodies regarding New Zealand's relatively large literacy achievement gap continued for such a long period of time? Our contention is that there is a straight forward answer to this question, which is that the New Zealand national literacy strategy is simply not working.

The remainder of this paper is divided into three sections. The first section presents findings in support of our claim that New Zealand's national literacy strategy has failed. The second section presents arguments and evidence in support of what we believe are the major factors responsible for the persistence of New Zealand's comparatively wide gap in literacy achievement, and for why the gap has not diminished over the past 15 years despite major efforts by the MoE to address the problem. The third section reviews research in support of what we maintain is the most effective strategy for reducing the literacy achievement gap: the use of differentiated instruction from the outset of formal schooling that takes into account

interactions between school entry reading-related skills (high vs. low literate cultural capital) and method of teaching reading (constructivist vs. explicit approaches).

EVIDENCE THAT NEW ZEALAND'S NATIONAL LITERACY STRATEGY HAS FAILED

For over 20 years New Zealand has consistently shown comparatively high levels of variability in the test scores from international surveys of reading achievement (Tunmer, Chapman, & Prochnow, 2003, 2004, 2006; Tunmer et al., 2008; Tunmer & Prochnow, 2009; Tunmer, Prochnow, Greaney, & Chapman, 2007). The high degree of variability in outcomes is somewhat unexpected for two reasons. First, New Zealand has a unified national education system with a relatively uniform approach to literacy instruction and intervention. Most aspects of literacy education are controlled centrally by the MoE, including the setting and monitoring of the national curriculum, the establishment of national reading and writing standards, the production of beginning reading materials and instructional guides for beginning teachers, and the funding and monitoring of two major intervention programmes for struggling readers, Reading Recovery and Resource Teachers: Literacy (Chamberlain, 2012). Consequently, compared with other English-speaking countries like the United States or Canada (that have semi-autonomous education systems at the state or provincial level), there is considerably less variation in the materials, reading methods, and instructional strategies used in regular classroom reading programmes and in nationally implemented intervention programs.

The second reason New Zealand's relatively large literacy achievement gap is rather surprising concerns Reading Recovery (RR), a nationally implemented early intervention programme developed by Clay (1985) to help children identified as making only limited progress in reading after a year of formal reading instruction (normally children whose reading progress falls in the lowest 15 to 20 percent of the enrolment cohort in any given school). The programme involves one-to-one pull-out instruction for 30 to 40 minutes per day for a period of 12 to 20 weeks by a specially trained RR teacher (Clay, 2005 a, b). The main goal of the programme is to accelerate students' reading achievement to the *average* level of their peers within a 20-week period (Chamberlain, 2012; Lee, 2011). Clay (1987) was very confident about the effectiveness of RR, claiming that it 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 economic causes] and all the children who have organically based problems but who can be taught to achieve independent learning status in reading and writing despite this" (p. 169). However, if the RR programme had been successful in attaining its goal of substantially reducing the number of children who develop ongoing reading difficulties, then the relatively large gap in reading performance consistently observed between good and poor readers since the 1991 IEA literacy study (Elley, 1992) should have steadily decreased after RR was introduced throughout the country in the late 1980s. This has not been the case.

In addressing the question of whether New Zealand's national literacy strategy has failed, we draw on data from two recently released reports that are readily accessible: the PIRLS 2011 report, which was released in December, 2012 (Mullis, Martin, Foy, & Drucker, 2012), and the latest annual monitoring report of Reading Recovery data in New Zealand (Lee, 2011).

PIRLS 2011

The PIRLS is the most recent test of reading achievement developed by the IEA and focuses on the achievement and literacy learning experiences of children from countries throughout the world in grades equivalent to Year 5 in New Zealand (children enter Year 1 in New Zealand on or soon after their 5th birthday). The PIRLS, which was developed by Mullis et al. (2003), is a 5-year cycle of assessments that was first administered in 2001, then in 2005/2006, and again in 2010/2011. It includes a state-of-the-art test of reading comprehension that was designed to assess two aspects of reading literacy: purposes of reading and processes of comprehension. For the 2001 assessment, scaling procedures based on Item Response Theory were used to establish an international scale for the reading comprehension test with a mean of 500 and a standard deviation of 100. To monitor increases or decreases in average reading achievement across assessment cycles, data from subsequent assessments were linked to this scale by making the 2001 scale centerpoint of 500 the reference point for the 2006 and 2011 studies (Mullis et al., 2012, p.36). The PIRLS also includes a series of questionnaires, given to principals, teachers, parents, and students, to obtain information on reading behaviours, reading attitudes, and home and school contexts for learning to read.

As shown in Table 1, the general pattern of results observed for New Zealand in earlier PIRLS assessments continued to be displayed in the PIRLS 2011 study, which focused on the achievement and reading experiences of children in 45 countries/regions. The mean achievement scores for individual countries ranged from a high of 571 (achieved by Hong Kong Special Administrative Region, SAR) to a low of 310 (the mean score for Morocco). Because there was a long tail in the distribution of the means of participating countries, 32 countries scored significantly higher than the centerpoint and 12 countries scored significantly lower. New Zealand was ranked 23rd with a mean score of 531.

Table 1
Means, standard deviations, and percentiles of the reading achievement scores for New Zealand as a function of PIRLS test cycle

Test Cycle	Mean Scale Score	Standard Deviation	Percentile Scores	
			5th Percentile	95th Percentile
PIRLS 2001	529	93	360	668
PIRLS 2006	532	87	374	664
PIRLS 2011	531	88	373	666

The means of 20 of the 45 participating countries were significantly higher than the New Zealand mean, and the means of 17 countries were significantly lower. This was the second time since New Zealand began participating in studies of reading achievement by the IEA that the number of countries that significantly outperformed New Zealand exceeded the number of countries that New Zealand significantly outperformed. The first time was in the PIRLS 2006 study, when 21 of the 45 participating countries scored significantly higher than New Zealand and 19 countries scored significantly lower.

Given the association of lower reading achievement with various economic factors (such as residing in low-income families, living in poor neighbourhoods, having parents with limited education and low levels of literacy, and attending schools in which literacy achievement is chronically low; Snow, Burns & Griffin, 1998) and New Zealand's relatively high level of economic development (ranked 30th in the world in GDP per capita by the World Bank), New Zealand would be expected to perform better than countries that are underdeveloped and/or have populations with large differences in material wealth, such as the lowest performing countries in the PIRLS 2011 study (Malta, Trinidad and Tobago, Azerbaijan, Iran, Colombia, United Arab Emirates, Saudi Arabia, Indonesia, Qatar, Oman, and Morocco). A more useful comparison would be to consider countries that are more similar to New Zealand in respect of economic development, language of instruction, linguistic homogeneity, and complexity of orthography. Six countries satisfied these criteria: Northern Ireland, the United States, Ireland, England, Canada, and Australia. Although the mean score for Australia did not differ significantly from New Zealand's mean score, the mean of each of the five remaining countries was significantly higher than the New Zealand mean. Overall, Northern Ireland was ranked fifth, the United States sixth, Ireland tenth, England eleventh, and Canada twelfth.

For countries that participated in earlier PIRLS assessments, it was possible to examine trends across the participants during the decade of 2001 to 2011 (Mullis et al., 2012, pp. 46-50). Overall, there were more increases than decreases in reading achievement. From 2001 to 2011, 48% of the countries had significant gains in reading achievement compared with 33% showing no change and 19% showing significant decreases. From 2006 to 2011, 46% of the countries achieved significantly higher mean scores compared with 29% showing no change and 25% showing significant decreases. Of the two English-speaking comparison countries for which trend data were available, both the United States and England achieved significantly higher scores in PIRLS 2011 than in PIRLS 2006. In contrast, for New Zealand there were no significant differences between the mean scores of any of the PIRLS assessment points (2001 vs. 2006, 2001 vs. 2011, or 2006 vs. 2011).

In view of the ongoing concerns expressed by New Zealand policy makers about the large inequities in outcomes in literacy education, the relative spread of New Zealand's reading achievement scores in the PIRLS 2011 assessment was of particular interest. The standard deviation of 88 and associated difference of 293 scale points between the 5th and 95th percentiles were exceeded by only seven of the 45 participating countries, a finding very similar to what was reported in the PIRLS 2001 and 2006 studies (see Table 1). Of the six English-speaking comparison countries, none had a standard deviation as large as New Zealand's. Presented in Table 2 are the mean reading achievement scores for New Zealand as a function of ethnicity and PIRLS test cycle (see Chamberlain & Caygill, 2012, p. 7, and earlier MoE PIRLS reports accessible at www.educationcounts.gov.nz). The differences in mean reading achievement scores favoring Pākehā/European over Māori students were 71, 69, and 70 for PIRLS 2001, 2006, and 2011, respectively, and the corresponding differences in scores favoring Pākehā/European over Pasifika students were 71, 73, and 85. The size of the reading achievement gap between Pākehā/European and Māori/Pasifika students has not decreased over the past decade and, if anything, has become slightly larger.

In addition to the high level of disparity between New Zealand's good and poor readers, international benchmarks based on the type of questions students were able to answer (from locating and reproducing explicitly stated information in text to integrating complex information from different parts of text) showed that in the PIRLS 2011 study, New Zealand had relatively large proportions of students performing at both the highest and lowest levels.

The percentage of New Zealand students performing above the advanced international benchmark (14%) was exceeded by only nine countries, whereas the percentage of students who failed to reach the low international benchmark (8%) was exceeded by only 13 countries. Of particular importance, there were no significant changes in the percentages of New Zealand students reaching the four international benchmarks (low, intermediate, high, and advanced) since the PIRLS 2001 or 2006 studies (Mullis et al., 2012, p. 70). In contrast, for other countries that participated in earlier PIRLS assessments (2001 and/or 2006), the trend data revealed that in general, there were more significant improvements in the percentages of students reaching the international benchmarks in 2011 than there were declines. For both the United States and England, there were significant improvements across the international benchmarks and no significant declines.

Table 2
Mean reading achievement scores for New Zealand as a function of ethnicity and PIRLS test cycle

	Ethnic Group			
Test Cycle	Pākehā/European	Asian	Māori	Pasifika
PIRLS 2001	552	540	481	481
PIRLS 2006	552	550	483	479
PIRLS 2011	558	542	488	473

In summary, despite the attempts by the New Zealand government over the past decade to reduce the relatively large disparity between good and poor readers, our examination of the PIRLS 2011 results has revealed that these efforts have largely failed. Virtually no changes in educational outcomes have occurred. The average reading achievement score for New Zealand in the PIRLS 2011 study was not significantly different from either the PIRLS 2001 or 2006 studies (see Table 1). As occurred in the PIRLS 2006 study, the number of countries that significantly outperformed New Zealand exceeded the number of countries that New Zealand significantly outperformed. Of the six English-speaking comparison countries, all but one significantly outperformed New Zealand. Trend data revealed that, although there were more increases than decreases in mean reading achievement scores across countries from 2001 to 2011, New Zealand showed no significant increases in reading performance. The standard deviation and range (between the 5^{th} and 95^{th} percentiles) for New Zealand's reading scores were almost unchanged from the PIRLS 2001 and 2006 studies and exceeded the values of most other countries, including those of the six English-speaking comparison countries. The large differences in reading achievement scores between Pākehā/European and Māori/Pasifika students have also not changed over the past decade (see Table 2). Despite general improvement across countries in the percentages of students reaching international benchmarks from 2001 to 2011, there were no significant changes from the PIRLS 2001 or 2006 results in either the relatively high percentage of New Zealand students who performed at the advanced international benchmark or the relatively high percentage of students who failed to reach the low international benchmark.

Reading Recovery 2011

The fact that no progress has been made in closing the literacy achievement gap raises questions about the effectiveness of RR, which has been operating on a national basis in New Zealand for over 25 years and has as its stated purpose bringing struggling readers to average levels of reading performance within 20 weeks. A possible explanation for the apparent failure of RR to have had a significant impact on reducing the literacy achievement gap is that the programme is of limited benefit to those struggling readers who need help the most. In a review of research on RR, Reynolds and Wheldall (2007) reported that the programme "has not demonstrated that it works for the students who are most at risk of failing to learn to read" (p. 213), leading them to conclude that "the success of the programme appears to be inversely related to the severity of the reading problem" (p. 219). The results of international studies of literacy achievement since 1991 have consistently shown that the low performing readers in New Zealand were likely to be Māori or Pasifika and/or from low-income backgrounds. Given Reynolds and Wheldall's claims, these students would be expected to benefit the least from RR. To test this prediction we examined data from the latest annual monitoring report of RR in New Zealand, which includes trend data for the past 10 years (Lee, 2011).

The annual monitoring report is based on two sources of information provided by schools: an end-of-year school report and individual student reports (Lee, 2011). The school report includes information on the number of students involved in RR and the number of hours and teachers allocated to RR for the year. The student reports include information on the demographic/background characteristics of the student, the amount of time spent in RR, outcome from RR, and entry and exit scores on three assessment tools: instructional text level, the Burt Word Reading Test, New Zealand Revision (Gilmore, Croft, & Reid, 1981), and the Writing Vocabulary Task (Clay, 2002). A running record of the child's oral reading behaviour as he or she reads a selected text provides the basis for assigning text level, which is the level of books that the child is able to read with 90% to 94% word recognition accuracy. The Burt Word Reading test is a measure of context-free word recognition in which the child is asked to read words of increasing difficulty. Blaiklock (1997) reported that the Burt test correlated highly with word recognition accuracy in connected text (r = .94) and reading comprehension ability (r = .85). In the Writing Vocabulary Task, the child is given a maximum of 10 minutes to write all the words that they know on a blank sheet of paper. A standard set of prompts is given, if necessary.

In 2011, 64.1% of state schools with 6-year-old populations offered RR, which served 75.2% of the total 6-year-old school population. Of the total 6-year-old population in state schools, 13.8% of the students entered RR in 2011 (approximately one in seven students). This percentage has remained fairly stable since 2002, ranging between 14% and 15% (see Figure 1 of Lee, 2011). Of the total 6-year-old population in schools offering RR, 18.4% of the students entered RR in 2011 (just under one in five students), and 25.0% of the students (one in four) were involved in RR at some point during the year (which included RR students carried over from 2010).

In New Zealand, schools are rated from a decile of 1 (low) to 10 (high) according to the socio-economic community the school serves. The 2011 monitoring report indicated that RR was more likely to be implemented in high decile (8 to 10) schools (71.4%) than in low decile (1 to 3) schools (56.4%). However, in low decile schools, 17.1% of students entered RR compared with 11.2% of students attending high decile schools. Further underscoring this

point, the average RR hours allocated per student in decile 1 schools was 51.7 hours compared with 42.9 hours per student in decile 10 schools. That is, the schools that were more likely to have students needing remedial assistance were less likely to offer RR, and when they did, the children placed in RR spent more time in the program.

A related finding is that Māori and Pasifika students were less likely to attend schools that offered RR than students in the total 6-year-old population. Access to RR for the total 6-year-old population was 75.2% compared with 69.7% for Māori students and 72.6% for Pasifika students (see Tables 2, 3, and 4 of Lee, 2011). That both Māori and Pasifika students and students in low decile schools were less likely to attend schools with RR was not surprising, as a disproportionate number of Māori and Pasifika students come from low-income backgrounds. The lower rate of implementation of RR in schools with Māori and Pasifika students (especially those with relatively large proportions of Māori or Pasifika students, such as schools in parts of the Auckland region) may be due, in part, to the significant levels of concern expressed by principals and teachers about the effectiveness of RR for these students and the lack of sufficient places for the large numbers of students in these schools needing extra support (Chapman, Greaney, & Tunmer, 2007; McDowall, Boyd, Hodgen, & van Vliet, 2005).

Although Māori and Pasifika students constituted 34.9% of the total 6-year-old student population in 2011, they made up 44.0% of the students involved in RR (which included RR students carried over from 2010), consistent with ongoing trends (percentages derived from data presented in Table 8 of Lee, 2011). Of the total 6-year-old population of Māori and Pasifika students, 23.6% were involved in RR (33.5% of the total 6-year-old population of Māori and Pasifika students in schools that offered RR), compared with 17.3% for New Zealand (NZ) European/Pākehā students (percentages derived from data presented in Tables 3, 4, and 8 of Lee, 2011). The higher participation rate for Māori and Pasifika students in RR indicates that these students were already tending to fall behind in reading after only one year of formal schooling, which invites the question of why this is so. This issue will be addressed in the next section of the paper.

Regarding RR outcomes in 2011, 81.6% of RR students were successfully discontinued, 11.8% were referred on for specialist help or long-term support, 5.4% left the school before completing the program, and 1.2% were responding but not able to be continued (see Table 10 of Lee, 2011). Over the past decade 11% to 13% of RR students (approximately one in eight) did not complete the programme but, instead, were referred on for specialist help (see Figure 3 of Lee, 2011). Māori and Pasifika students were less likely than NZ European/Pākehā students to have been successfully discontinued from RR (76.1%, 80.9%, and 84.6 for Māori, Pasifika, and NZ European/Pākehā students, respectively) , and more likely to have been referred on for specialist help (13.8%, 12.2%, and 11.2% for Māori, Pasifika, and NZ European/Pākehā students, respectively), a pattern that has been observed in RR outcome data for the past 10 years. Of the total number of students who were referred on, almost half (49.1%) were Māori or Pasifika (percentage derived from data presented in Tables 10 and 13 of Lee, 2011).

A similar pattern was observed for students attending low decile schools. For deciles 1 to 3, 76.6% of students were successfully discontinued compared with 86.0% of students in deciles 8 to 10, whereas 14.4% of students in deciles 1 to 3 were referred on compared with 9.2% of students in deciles 8 to 10 (see Figure 4 of Lee, 2011). Regarding the latter, 15.2% of students in decile 1 schools were referred on, which was more than double the rate of 7.2% of

students in decile 10 schools who were referred on (see Appendix 1 of Lee, 2011). The results further indicated that, compared with students who were successfully discontinued from RR, students who were referred on for specialist help had been given, on average, more 30-min sessions (89.2 vs. 72.5) over a longer period of time (22.8 weeks vs. 17.8 weeks) (see Table 15 of Lee, 2011). In summary, Māori and Pasifika students and students from low decile schools (largely the same groups) were less likely to have been successfully discontinued from RR and more likely to have been referred on for specialist help. In addition, the referred on students had failed to respond adequately to RR despite having received extra lessons and more time in the program.

In the final section of the Lee (2011) report, students' gains from participating in RR are presented in the form of relative frequency distributions for each of the three assessment measures (instructional text level, Burt Word Reading, and Writing Vocabulary) on entry to and on exit from RR for successfully discontinued and referred on students. Three particularly striking patterns emerged from the results. First, students who were referred on consistently had much lower mean (or median) scores on the assessment measures on entry to RR than students who were successfully discontinued. For instructional text level, the median entry score was 5 for successfully discontinued students compared with 2 for referred on students. For Burt Word Reading, the mean entry score was approximately 13 for successfully discontinued students compared with 6 for referred on students. For Writing Vocabulary, the mean entry score was approximately 23 for the successfully discontinued students compared with 10 for the referred on students. In general, the entry scores of successfully discontinued students were much higher than those of the referred on students, which was also true for the exit scores: 18 versus 13 for instructional text level; 29 versus 17 for Burt Word Reading, and 57 versus 32 for Writing Vocabulary (see Figures 7, 8, 9, 10, 11, and 12 of Lee, 2011). Students who entered RR with relatively high scores on the assessment measures were clearly much more likely to benefit from RR than students with relatively low scores. These findings provide strong evidence in support of Reynolds and Wheldall's (2007) claim that RR generally does not work well for students who are most at risk of failing to learn to read. Given that the referred on students were more likely to have been Māori or Pasifika and/or from low-income backgrounds, the results also provide an explanation for the failure of RR to have had a significant impact on reducing the literacy achievement gap in New Zealand.

The second striking feature of the data on students' gains from RR was the relatively high level of variability in the assessment scores on entry to and on exit from RR. The entry and exit scores of successfully discontinued students for the Burt Word Reading test and the Writing Vocabulary Task overlapped to such an extent that some of the students had entry scores that exceeded the mean of the exit scores (see Figures 9 and 11 of Lee, 2011). This highly atypical pattern most likely reflects the fact that in New Zealand RR is made available to the lowest 15 to 20 percent of students compared with their cohort *in an individual school*, even if the school has a high decile rating. However, 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, Mullis et al., 2012; Tunmer et al., 2008; see discussion in the following section), students who qualify for RR in high decile schools are more likely to enter (and exit) the programme with higher scores on the assessment measures than students in low decile schools, which would account for the high level of variability in the assessment measures.

One of the key recommendations of the Literacy Taskforce (described in the preceding section) was that "in order to make the best use of available resources, Reading Recovery should be targeted to the children with the greatest need, particularly those in lower decile schools" (Ministry of Education, 1999b, p.21). However, this recommendation to restrict the availability of RR to the lowest performing 6-year-old students throughout the entire school system was rejected by the MoE, a decision that may have contributed to the relatively high false positive rate of approximately 30% of the children selected for participation in RR, as reported by Center, Wheldall, Freeman, Outhred, and McNaught (1995). That is, a large proportion of students selected for RR would have caught up with their peers even if they had not received RR instruction. Contributing further to the tendency to include in RR students who have relatively mild reading problems (that are often spontaneously overcome through regular classroom instruction without intervention) is evidence that a significant number of the lowest performing 6-year-olds are excluded from RR because they are considered not ready or less likely to benefit from the programme than other students, or are withdrawn early from the programme because they failed to make expected rates of progress (Chapman et al., 2007; Church, 2005; McDowall et al., 2005).

The third striking feature of the data on entry and exit scores concerns the distributional pattern for instructional text level (see Figure 7 of Lee, 2011). Unlike the distributions for the Burt Word Reading Test and Writing Vocabulary Task, the distribution for instructional text level on exit from RR for successfully discontinued students was non-overlapping with entry text levels, positively skewed, and truncated at text level 16 (the recommended minimal level for discontinuation from RR). This pattern is consistent with other New Zealand studies reporting marked discrepancies between RR teachers' and classroom teachers' assessments of text level at discontinuation (Chapman et al., 2001; Glynn, Crooks, Bethune, Ballard, & Smith, 1989). Independent and research-validated measures of reading performance (such as the Burt Word Reading test, which correlates highly with word recognition accuracy in connected text) have shown that classroom teachers' assessments of text level are more accurate. The associated reading ages for text level and independent measures such as the Burt test are more closely aligned for classroom teachers than for RR teachers. Text level as determined by running records appears to be an unreliable measure that yields inflated estimates of reading achievement (Blaiklock, 2004; Elbaum, Vaughn, Hughes, & Moody, 2000; Hiebert, 1994; Tunmer & Chapman, 2003).

A major aim of RR is to help struggling readers to develop a "self-extending system" of reading and writing strategies so that they can learn effectively in the regular classroom without additional support (Clay, 1991). However, the annual monitoring reports do not provide data on the long-term benefits of RR, and there is limited New Zealand research showing that the gains from participating in RR are sustained over time. What little evidence there is regarding maintenance effects indicates that the gains made by students successfully discontinued from RR largely disappear within 2 years following completion of the programme (Chapman et al., 2001; Glynn et al., 1989). In a more recent study Limbrick and Jesson (2010) examined the standardized achievement scores of 373 students 2, 3, or 4 years after they had been successfully discontinued from RR. Results indicated that only one third (34%) of the students achieved stanine 5 or above on the Supplementary Test of Achievement in Reading (STAR: Elley, 2001) or the Progressive Achievement Test of Reading Comprehension (PAT: New Zealand Council for Educational Research, 2008). The remaining two-thirds achieved at stanine levels below average (approximately 29% at stanine 4, 22% at stanine 3, and 15% at stanines 1 and 2). Consistent with these findings are data from the most recent annual report for Resource Teachers: Literacy (RT:Lit), specialist

teachers who assist older students with persistent literacy learning difficulties (Lee, 2012). Data on prior involvement in RR indicated that nearly one third (31.6%) of RT:Lit students in 2011 had previously received RR, and that of these students, just over a third (34.9%) had been successfully discontinued from RR, which is consistent with ongoing trends (see Tables 4 and 5 of Lee, 2012).

In summary, New Zealand research has revealed that RR is less accessible to Māori and Pasifika students and students from low-income backgrounds (most likely because of RR's perceived ineffectiveness for these students) and that a significant number of the lowest performing 6-year-olds are excluded from RR (because they are considered not ready or less likely to benefit from the programme or are withdrawn early from RR because they failed to make expected rates of progress). Māori and Pasifika students and students from low-income backgrounds are less likely to be successfully discontinued from RR and more likely to be referred on for specialist help (despite having received extra lessons and more time in the program). In support of Reynolds and Wheldall's (2007) claim that RR generally does not work well for students who are most at risk of failing to learn to read, students who enter RR with relatively high scores on the assessment measures of RR (most likely students from high decile schools) are much more likely to benefit from RR than students with relatively low scores (most likely students from low decile schools). Finally, research indicates that positive maintenance effects for the majority of successfully discontinued RR students are modest or non-existent. For these reasons RR has had little or no impact on reducing New Zealand's relatively large literacy achievement gap.

WHY HAS NEW ZEALAND'S NATIONAL LITERACY STRATEGY FAILED?

Three interrelated factors have contributed to the failure of New Zealand's national literacy strategy: a constructivist orientation toward literacy education, the failure to respond adequately to differences in literate cultural capital at school entry, and restrictive policies regarding the first year of literacy teaching.

Pedagogical constructivism in literacy education

For the past 25 years New Zealand has followed a predominantly constructivist approach to literacy education that assumes that learning to read is essentially like learning to speak, where both abilities are thought to develop "naturally" (Smith & Elley, 1994, p.81). In a review of Australian and New Zealand reading research, Wilkinson et al. (2000) noted that "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 favor" (p.12). Smith and Elley (1994), two leading proponents of the constructivist approach to teaching reading in New Zealand, claimed that "children learn to read themselves; direct teaching plays only a minor role" (p.87). In this approach literacy learning is largely seen as the by-product of active mental engagement with little or no explicit, systematic teaching of phonemic awareness (the ability to reflect on and manipulate the phonemic segments of spoken words) and alphabetic coding skills (the ability to translate letters and letter patterns into phonological forms). Smith and Elley (1994) argued that teaching beginning readers orthographic patterns "is a difficult, unnecessary and largely fruitless activity, creating distorted ideas about the nature and purpose of reading" (p.143). Explicit instruction in word-level skills and strategies is therefore downplayed or discouraged. Word analysis activities, if any, arise primarily from the child's responses during text reading and focus mainly on initial letter sounds.

Underpinning the constructivist approach to literacy teaching is the "multiple cues" theory of reading (Tracey & Morrow, 2006), which is also referred to as the "searchlights" model of reading (Rose, 2006). According to this view, skilled reading is a process in which minimal word-level information is used to confirm predictions about the upcoming words of text based on multiple sources of information. Clay (1991) stated that "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 the 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). Smith and Elley (1994) expressed a very similar view, claiming that because language follows a predictable pattern (which is a false assumption; see Pinker, 1994; Tunmer & Nicholson, 2011), children "learn to read with minimal input from the text, predicting and confirming and making sense as they go" (p. 142).

Unlike skilled readers, according to this view, poor and beginning readers are less able to make use of contextual redundancy in ongoing sentence processing. Reading acquisition is therefore seen largely as a process in which children learn to use multiple cues in identifying words in text, with text-based cues (i.e., picture cues, sentence context cues, preceding passage context, prior knowledge activated by the text) being used to generate predictions about the text yet to be encountered and letter-sound information generally being used for confirmation and self-correction. As Clay (1998) argued, 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; and special knowledge from past literary 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).

These theoretical assumptions about the nature of skilled reading and reading acquisition and the role of pedagogical constructivism in literacy education emerged in New Zealand during the 1980s. As Connelly, Johnston, and Thompson (2001) noted, the shift from emphasis on words in teaching reading in New Zealand to an emphasis on the story and book "has become more prevalent in the last twenty years and there has been increasing concern that children are able to predict reading responses from story and sentence context" (p.433). This perspective was adopted and strongly promoted by the MoE through its various publications. *Reading in Junior Classes* (Ministry of Education, 1991), the guidebook used by beginning reading teachers in New Zealand until it was replaced in 2003, explicitly stated that, "It is better that children predict meaning from other cues at the outset and use their knowledge of letters and sounds for confirmation" (p.48). Similarly, *The Learner as Reader* (Ministry of Education, 1996) stated that the first strategy children should be encouraged to use when confronted with an unknown word in text is to "try reading from the beginning of the sentence again and think what would fit" (p.50). In *Reading and Beyond* (Ministry of Education, 1997), the introduction to the *Ready to Read* series used in New Zealand schools, reading is described as "a constantly repeated process of sampling, predicting, checking, confirming, and self-correcting" (p.7). Readers "predict or anticipate the meaning of unknown text … [and] check these predictions against what they have read and what they know of the world, and either confirm or self-correct predictions by rereading, reading on, or referring to other cues" (p.7).

Reading in Junior Classes was replaced by the guidebook, *Effective Literacy Practice in Years 1 to 4* (Ministry of Education, 2003a), copies of which were distributed to every

teacher of Years 1 to 4 students throughout the country. *Effective Literacy Practice* was the cornerstone of the MoE's literacy strategy during this period (2003-2006) and the key resource for a large-scale, in-service professional development programme implemented by the MoE. The new guidebook continued to endorse strongly the multiple cues/searchlights theory of reading, stating that "fluent readers … draw on their prior knowledge and use all available sources of information simultaneously and usually unconsciously" (p.30), and that "in skilled reading, predictions are usually checked swiftly and automatically" (p.130). Based on these (invalid) assumptions about skilled reading, *Effective Literacy Practice* stated that teachers need to show beginning readers how to "cross-check *predictions* to ensure that they make sense and fit with other information already processed" and that "for beginning readers, cross-checking usually involves checking that their *prediction* of an individual word fits and makes sense" (p.130, emphasis added).

There is clear evidence from classroom observations and teacher surveys that the advice presented in MoE publications has been widely implemented in New Zealand classrooms. In a study of the prompts teachers prefer to use when assisting beginning readers to identify unfamiliar words in text, Greaney (2001) found that text-based prompts were much more likely to be selected over prompts that encouraged the use of word-based strategies. Although *Effective Literacy Practice* acknowledges the importance of phonemic awareness and alphabetic coding skills in learning to read, explicit, systematic instruction in these skills is discouraged. Instead, it recommends that these skills should largely be taught through reading and writing activities, as described in *Sound Sense: Phonics and Phonological Awareness* (Ministry of Education, 2003c), where it is stated that, "children are more likely to make connections between phonics and their reading and writing of texts if they are engaged and involved in making discoveries for themselves" (p.7).

According to multiple cue theorists, focusing too much attention on the development of word-level skills and strategies may actually contribute to reading failure by diverting the child's attention away from what are considered to be more productive strategies. Clay (2005a), for example, warned that "undue attention to the detail of letters … can block the child's ability to use his language knowledge and the meaning of text, as part of his information base for decision making" (p.25). Clay (2005b) explicitly stated that if the child has a bias toward using mainly letters to identify unknown words in text, "the teacher's prompts will be directed towards the message and the language structure" (p.112).
This approach to teaching reading is reflected in the use of the running record, the primary assessment tool used by both RR and regular classroom teachers in New Zealand. A running record is a copy of a read passage on which the student's oral reading errors are recorded. The procedure for analysing the errors was derived from miscue analysis, which was introduced in the 1960s by Ken Goodman (1967), a leading proponent of the multiple cues view of learning to read. In an example of the use of running records from Clay (2000), a child was asked to read an illustrated story entitled *The Bicycle*, the text for which is given below:

The clown got on
and the lady got on
and the boy got on
and the girl got on
and the bear got on
and the bicycle got …
squashed.

The child incorrectly read *lake* for *lady*, *box* for *boy*, *bil* for *bicycle*, and *square* for *squashed*. According to the analysis of these errors, because the child already appears to be using visual cues (i.e., initial letters) and structural cues (i.e., syntax), the teacher should encourage the child to make greater use of meaning cues when attempting to identify unfamiliar words. However, this would be a very ineffective learning strategy to use, as research has firmly established that text-based cues should only be used to supplement word-level information, not to substitute for it (Snow & Juel, 2005; Tunmer & Nicholson, 2011).

The recently introduced national literacy standards (Ministry of Education, 2009) described previously further demonstrate the extent to which the multiple cues approach to teaching reading has become entrenched in New Zealand literacy education. As noted in our description, the standards include examples of reading behaviours that teachers would be expected to observe in students meeting the standards at each Year level. For the reading standard after one year at school, examples are provided of behaviours that may be exhibited by students who successfully identify unfamiliar words in text. Emphasis is placed on the use of multiple sources of information, especially text-based cues. For example, in identifying the unfamiliar word *rode* in the sentence, "My great-grandma rode a horse," the standards state that "the student may use context (including the illustration), any prior knowledge about horse riding, and their knowledge of language structures and letter-sound relationships" (p.20). In an example in which the student "works out" the unfamiliar words *shady*, *bubbly*, and *plenty*, the standards state that, "The student uses their accumulated understanding of the story, their knowledge of language structures, and their knowledge that 'y' at the end of a word can have an 'ee' sound" (p.20). In both examples the order in which the different sources of information is mentioned is consistent with Clay's (1998) view that letter-sound information should only be used "in the last resort" (p.9).

Evidence against the multiple cues/searchlights model of reading

Our contention is that the MoE's rigid adherence to the multiple cues/searchlights model of reading as the theoretical basis for its approach to literacy teaching and assessment has contributed greatly to the continuing inability to reduce the literacy achievement gap. Following thorough reviews of the scientific literature on learning to read, countries throughout the world have abandoned the searchlights model of reading. In the United Kingdom, for example, the searchlights model was rejected in favor of the Simple View of Reading (SVR) model of individual differences in reading comprehension performance (Gough & Tunmer, 1986; Hoover & Gough, 1990; Tunmer & Chapman, 2012b; Tunmer & Greaney, 2010; Tunmer & Hoover, 1993; Vellutino, Tunmer, Jaccard, & Chen, 2007). The SVR model provided the conceptual framework underlying the wide range of recommendations included in the Rose (2006) report, and has been adopted as the theoretical basis of the revised national curricular advice to all schools in England regarding the teaching of decoding and comprehension skills (Department of Education, 2010).

The scientific community has also firmly rejected the multiple cues/searchlights model of reading. The major shortcoming of the instructional philosophy espoused by multiple cue theorists is that it stresses the importance of using information from many sources in identifying unfamiliar words in text without recognizing that skills and strategies involving phonological information are of primary importance in beginning literacy development. As Pressley (2006) pointed out, "the scientific evidence is simply overwhelming that letter-sound cues are more important in recognizing words … than either semantic or syntactic cues" (p.21), and that "teaching children to decode by giving primacy to semantic-contextual and

syntactic-contextual cues over graphemic-phonemic cues is equivalent to teaching them to read the way weak readers read!" (p.164). One of the major distinguishing characteristics of struggling readers is their tendency to rely heavily on sentence context cues to compensate for their deficient alphabetic coding skills (Stanovich, 1986).

Supporting these claims is research carried out in New Zealand by Tunmer and Chapman (2002) examining the relation of beginning readers' reported strategies for identifying unknown words in text to future reading achievement. Five-year-old (Year 1) beginning readers participating in a 3-year longitudinal study were divided into two groups according to their responses to the following question: "When you are reading on your own and come across a word that you don't know, what do you do to try to figure out what the word is?" The majority of children (52.1%) reported using word-based strategies (e.g., sound it out, think of the sounds, say the letters, do the sounds of it, make the sounds, hear all the letters, listen to what the letters are, you try and get the letters right, you hear the letters, say out the sounds in the word, sound it out – dad says so) rather than text-based strategies (e.g., guess, think, guess what the word is, read it over again, read on, have a look at the picture, keep on going – then go back and see what the word is, I leave it, think about the word, try to guess what it is).

From these results Tunmer and Chapman concluded, as did Liberman and Liberman (1992), that the instructional approach to teaching reading followed in the classroom (i.e., the constructivist, multiple cues approach) is not necessarily reflected in the word identification strategies that the majority of children actually use in learning to read. Of greater importance, the results further showed that the Year 1 beginning readers who reported using word-based strategies strongly outperformed the children who reported using text-based strategies on all reading and reading-related measures taken in the middle of Year 3. Moreover, these children were six times less likely to enter RR in Year 2 than the children who relied on text-based strategies in Year 1 (6% vs. 37%). Also, Pākehā/European children were three times more likely to report using word-based strategies in Year 1 than Māori/Pasifika children (62% vs. 21%).

Research on how children learn to read indicates that achievement in reading comprehension performance depends on the ability to recognize the words of text accurately and quickly. For progress to occur in learning to read, the beginning reader must acquire the ability to translate letters and letter patterns into phonological forms (Ehri, 2005; Snow & Juel, 2005; Tunmer & Nicholson, 2011). Making use of letter-sound relationships provides the basis for constructing the detailed orthographic representations required for the automatization of word recognition (or what Ehri, 2005, calls *sight word* knowledge), thus freeing up cognitive resources for allocation to sentence comprehension and text integration processes (Pressley, 2006). To discover mappings between spelling patterns and sound patterns, children must also be able to segment spoken words into subcomponents. Children who experience ongoing difficulties in detecting phonemic sequences in words (i.e., phonemic awareness) will not be able to fully grasp the alphabetic principle and discover spelling-to-sound relationships (Shankweiler & Fowler, 2004). As the reading attempts of beginning readers with a firm understanding of the alphabetic principle become more successful, they will begin making greater independent use of letter-sound information (possibly supplemented with sentence context cues) to identify unfamiliar words in text. Phonologically decoding words a few times ultimately cements the orthographic representations of the words in lexical memory from which additional spelling-sound relationships can be induced without explicit instruction (Snow & Juel, 2005; Tunmer & Nicholson, 2011).

Although children must rely increasingly on induction to acquire the spelling-sound relationships necessary for learning to read, explicit phonics instruction plays an important role in helping to "kick-start" the process by which beginning readers acquire untaught spelling-sound relationships through implicit learning. Venezky (1999) argued that "phonics is a means to an end, not an end itself" (p.231). Accordingly, phonics has been defined as "an approach to, or type of, reading instruction that is intended to promote the discovery of the alphabetic principle, the correspondences between phonemes and graphemes, and phonological decoding" (Scarborough & Brady, 2002, p.20).

One of the main functions of phonics instruction is to provide beginning readers with a process for generating approximate phonological representations of unknown words that gets them close enough to the correct phonological form that, with context, the correct identification can be made. Children learn to use their knowledge of spelling-to-sound relationships acquired through phonics instruction to produce approximate phonological representations, or partial decodings, for unknown words, especially those containing irregular, polyphonic (e.g., *ear* as in *bear* and *hear*), or orthographically complex spelling patterns. The phonological representations provide the basis for generating alternative pronunciations of target words until one is produced that matches a word in the child's lexical memory and makes sense in the context in which it appears. Additional spelling-sound relationships, especially context-sensitive patterns (i.e., those that depend on position-specific constraints or the presence of "marker" letters), can then be induced from the stored orthographic representations of words that have been correctly identified (Tunmer & Chapman, 2012a). For children encountering difficulty in developing the ability to perceive intuitively the redundant patterns and connections between speech and print, explicit instruction in alphabetic coding skills is likely to be crucial (Ryder, Tunmer, & Greaney, 2008).

In support of these claims is a large body of research indicating that explicit, systematic instruction in the code relating spellings to pronunciations positively influences reading achievement, especially during the early stages of learning to read (Brady, 2011; Hattie, 2009; National Reading Panel, 2000; Snow & Juel, 2005; Tunmer & Arrow, 2013). From an examination of findings from a wide range of sources that included studies of reading development, specific instructional practices, and effective teachers and schools, Snow and Juel (2005) concluded that explicit attention to alphabetic coding skills in early reading instruction is helpful for all children and crucial for some.

The strongest evidence supporting the effectiveness of systematic phonics instruction over non-systematic or no phonics instruction (e.g., constructivist approaches) comes from large-scale meta-analyses. The National Reading Panel (2000) reported the results of 66 treatment-control comparisons from 38 studies that satisfied strict criteria. Results indicated that the overall mean effect size at the end of training was .41; that the effectiveness of systematic phonics instruction was greater among kindergarteners and first graders (d = .55) than second through sixth grades (d = .27); that first grade children at risk benefitted more from phonics instruction (d = .74) than normally achieving first graders (d = .48); and that systematic phonics instruction had larger effects for children from low SES backgrounds (d = .66) than for children from middle SES backgrounds (d = .44).

More recently, Hattie (2009) reported even stronger effects for phonics instruction. He summarized 14 meta-analyses involving 12,000 students and found that phonics had a

significant effect on learning to read, with a mean effect size of .60. In contrast, he summarized four meta-analyses of constructivist, whole-language teaching involving 630 students and found almost no effect. The mean effect size was only .06.

RR, which was designed to complement regular classroom literacy instruction in New Zealand, is also based on the multiple cues theory of reading. In RR lessons particular emphasis is placed on developing within struggling readers a self-extending system of reading strategies that involves the flexible use of multiple cues (syntactic, semantic, visual, graphophonic) to detect and correct errors while reading text (Clay, 2005a,b). Struggling readers are encouraged to engage in the strategies of predicting, cross-checking, and confirming while attempting to identify unfamiliar words in text. However, as noted from our examination of data from the latest RR annual monitoring report (Lee, 2011), RR is of limited benefit to those struggling readers who need help the most. RR appears to be beneficial for some struggling readers but not others, as indicated by the high percentage (around 15% in New Zealand but up to 30% elsewhere) of RR students who do not complete the programme but, instead, are referred on by their RR teacher for further assessment and possible remedial assistance (Chapman et al., 2007; Elbaum, Vaughn, Hughes, & Moody, 2000; Lee, 2011).

The differential effectiveness of RR can be explained in terms of contemporary theories of reading development that assume that learning to read is a cognitive-developmental process that takes place over time and may break down at different points (Tunmer & Nicholson, 2011). As a consequence of limited knowledge of print at the outset of learning to read and/or a developmental delay in acquiring phonological awareness skills, many struggling readers take longer than usual to acquire the self-improving alphabetic coding skills necessary for achieving progress in learning to read. For these children, the heavy emphasis on text reading in RR lessons provides them with additional opportunities to apply their developing alphabetic coding skills to word identification. As their reading attempts become more successful, these delayed readers will begin making greater independent use of letter-sound information to identify unfamiliar words from which additional spelling-sound relationships can be induced without explicit instruction. The extra practice in reading provided in RR, which focuses on text reading, is therefore likely to be beneficial in helping these delayed readers catch up with their peers.

A large proportion of struggling readers, however – probably a majority of the students with literacy learning difficulties in low decile schools – operate at even lower developmental phases of word learning, which Ehri (2005) described as the pre-alphabetic and partial-alphabetic phases. Delayed readers who are still in these phases of reading development typically have limited or severely limited phonemic awareness and alphabetic coding skills. As noted previously, children who experience ongoing difficulties in detecting phonemic sequences in words will not be able to fully grasp the alphabetic principle and discover spelling-to-sound relationships. For these children, more intensive and systematic instruction in phonemic awareness and phonemically-based decoding skills is likely to be required than what is typically provided in RR lessons.

Evidence in support of this claim comes from two studies. Iversen and Tunmer (1993) reported that the effectiveness of RR could be improved considerably by incorporating into the programme more intensive and explicit instruction in phonological awareness and the use of letter-sound relationships (especially orthographic analogies), in combination with strategy training on how and when to use this knowledge to identify words while reading text and to

spell words while writing messages. Chapman et al. (2001) found in a longitudinal study of RR in New Zealand that the students who failed to achieve significant progress or maintain the gains made in the programme had limited or severely limited phonemic awareness and phonemically-based decoding skills at the beginning of the program, during the year preceding entry into the program, and during the year following participation in the program. The students who received the most benefit from RR were more advanced in phonological skills at the beginning of the programme than students who derived little or no benefit from the program, and progress in learning to read following participation in RR was strongly related to phonological skills at discontinuation from the program. Similar findings have been reported in studies in Australia (Center et al., 1995) and the United States (Iversen & Tunmer, 1993). These finding are consistent with the widely held view among reading scientists that the primary phenotypic manifestation of developmental reading problems is poor context free word recognition ability and associated phonological processing difficulties (Tunmer & Greaney, 2010; Vellutino & Fletcher, 2005; Vellutino, Fletcher, Snowling, & Scanlon, 2004).

Literate cultural capital, Matthew effects in reading achievement, and PIRLS 2011

As noted previously, for the past 20 years New Zealand has consistently shown comparatively high levels of variability in the test scores from international surveys of reading achievement. We have argued that New Zealand's relatively wide spread of scores is largely the result of Matthew (rich-get-richer and poor-get-poorer) effects (Stanovich, 1986), triggered by a predominantly constructivist, multiple cues approach to reading instruction and intervention that fails to respond adequately to differences in literate cultural capital at school entry (Tunmer, Chapman, & Prochnow, 2003, 2004, 2006; Tunmer et al., 2008; Tunmer & Prochnow, 2009; Tunmer et al., 2007). Literate cultural capital is a generic term referring to literacy-related knowledge and abilities at school entry that are an outgrowth of activities in the home environment that support early literacy development (Tunmer & Nicholson, 2011).

Examples of literate cultural capital include oral language ability (promoted by verbal interaction in the home); familiarity with "book" or "decontextualized" language and basic understanding of concepts and conventions of printed language (promoted by adult storybook reading and talking about story content); knowledge of letter names and sounds (developed by exposure to ABC books and games); ability to produce preconventional spellings of words (developed by manipulating movable letters to form "invented" spellings; e.g., writing *color* as KLR, or *fairy* as FRE); sensitivity to the subcomponents of spoken words, or phonological awareness (promoted by playing rhyming and sound analysis games, being read books that increase phonological sensitivity, encouraging the use of invented spellings, and exposure to alphabet materials and games; e.g., pig Latin, I spy, nursery rhymes, Dr. Seuss books, ABC books); and sensitivity to the semantic and syntactic constraints of sentence contexts, or syntactic awareness (promoted by verbal interaction in the home, playing language games, engaging in linguistic humor, and being read storybooks). Tunmer and Nicholson (2011) reviewed theory and research on how these school-entry competencies contribute to early literacy development.

Research carried out in New Zealand and elsewhere indicates that children enter school with large individual differences in the experiences and competencies (i.e., literate cultural capital) essential for acquiring literacy, and that children who possess higher levels of literate cultural capital at the beginning of school profit more from literacy instruction, learn to read sooner, and read better than children who do not (Gilmore, 1998; Nicholson, 2003; Tunmer et al., 2003; Whitehurst & Lonigan, 2001). Supporting this claim is a large body of research

showing substantial predictive relationships between preschool measures of reading-related skills and later reading achievement (Tunmer & Nicholson, 2011). In a 7-year longitudinal study of literacy development in New Zealand, Tunmer et al. (2006) found that a composite measure of literate cultural capital (comprising two measures of phonological awareness, two measures of syntactic awareness, letter-name knowledge, and receptive vocabulary) at the start of school (when the mean age of the children was 5 years, 1 month) accounted for almost 50% of the variance in Year 7 reading comprehension performance after the effects of all other school-entry variables were controlled (SES, ethnicity, verbal working memory). For the children in the bottom quartile of the literate cultural capital scores at school entry, all were at least one year behind in reading age in Year 7. The average deficit was 2 years, 4 months.

The results further indicated that children from low-income and/or culturally diverse backgrounds had considerably less literate cultural capital when they arrived at school than did children from more advantaged backgrounds, similar to what others have reported (Goldenberg, 2001; Nicholson, 1997, 2003; Phillips & Lonigan, 2005; Snow et al., 1998). Home literacy environment appears to be the major contributing factor to social and cultural differences in entry-level pre-reading skills (Hart & Risley, 1995; Korat, Klein, & Segal-Drori, 2007; Nash, 1993, 1997; Nicholson, 1999; Snow et al., 1998). Although many children from low-income backgrounds do not struggle to learn to read, and a significant number from middle-class backgrounds do, the odds are generally stacked against economically disadvantaged children because they live in families that are under varying degrees of financial and social stresses; for example, unemployment, single parent households, large families, and poor living conditions (Crooks & Caygill, 1999; Nicholson, 1997, 1998). As a consequence, fewer resources are available for books, study materials, learning aids, and private study areas, and less adult time is available for engaging children in important literacy-related activities prior to school entry. Children from low-income backgrounds are therefore particularly susceptible to early reading difficulties because they often have not had the preschool exposure to the kinds of language play activities and early literacy experiences that promote the development of crucial literacy-related skills.

Children who do not possess sufficient levels of essential reading-related skills at the outset of formal reading instruction (and who are not provided with supplementary instruction to develop these competencies, especially phonological awareness) will be forced to rely increasingly on ineffective word identification strategies such as using picture cues, partial visual cues, and contextual guessing, the continued use of which inevitably leads to literacy learning difficulties (Pressley, 2006). Because of their ineffective word identification skills, these children not only receive less practice in reading but soon begin to confront materials that are too difficult for them, which results in avoidance of reading, inattentive behaviours, low expectations of success, and withdrawal from literacy learning tasks (i.e., negative Matthew effects). As a consequence, they are prevented from taking advantage of the reciprocally facilitating relationships between reading achievement and other aspects of development, which are referred to as positive Matthew effects. These developmental spinoffs include vocabulary growth, ability to comprehend more syntactically complex sentences, development of richer and more elaborate knowledge bases, and greater practice opportunities for building fluency and facilitating implicit learning of letter-sound patterns, all of which promote further growth in reading by enabling children to cope with more difficult materials. Reliance on counterproductive word identification strategies is exacerbated by the multiple cues/searchlights approach to teaching reading because emphasis is placed on encouraging beginning readers to use text-based cues (i.e., semantic-contextual

and syntactic-contextual cues) with insufficient attention being given to the development of phonemically-based word-level skills and strategies, especially among children who are less able to discover letter-sound patterns as a by-product of more general reading (Torgesen, 2004).

Findings reported in studies by Tunmer et al. (2003, 2004) and Tunmer et al. (2008) support the hypothesis that the constructivist, multiple cues approach to teaching reading in New Zealand fails to respond adequately to differences in literate cultural capital at school entry, which in turn triggers negative Matthew effects in reading. Tunmer et al. (2003) found that incorporating into Year 1 literacy education programmes in New Zealand supplementary materials and procedures designed to help children develop awareness of sound sequences in spoken words and make greater use of letter-sound patterns in identifying unfamiliar words produced significantly greater gains in reading achievement than the standard constructivist approach to teaching reading, especially for children from low-income, culturally-diverse backgrounds. Tunmer et al. (2004) analysed data from the PIRLS 2001 study (Mullis et al., 2003) and found that for specified differences in literate cultural capital possessed by children at school entry (as assessed by the Early Home Literacy Activities Index, the Index of Parents' Attitudes Toward Reading, and the Index of Home Educational Resources), New Zealand's "one size fits all" constructivist approach to literacy education was consistently associated with much larger differences in future reading achievement than most other participating countries. Tunmer et al. (2008) analysed data from the PIRLS 2006 study (Mullis et al., 2007) and obtained a very similar pattern of results.

The persistence of this pattern is demonstrated in the results from the PIRLS 2011 study (Mullis et al., 2012), which included questionnaires relating to home environment support for reading achievement. Measures from the study that are most closely linked to the literate cultural capital construct described previously include the Early Literacy Activities (ELA) scale and the Parents Like Reading (PLR) scale, both of which are expanded versions of scales used in earlier PIRLS studies. The ELA scale was based on parents' responses concerning the frequency of literacy-related activities with which they engaged their children prior to school entry. The activities included reading books, telling stories, singing songs, playing with alphabet toys, talking about things done, talking about things read, playing word games, writing letters or words, and reading aloud signs and labels. Children were assigned to one of three categories on the ELA scale according to parents' combined responses relating to the frequency with which they engaged their child in each of the nine activities. The mean percentage of students in each category across all 45 countries was 37% (*often*), 60% (*sometimes*), and 3% (*never or almost never*). Among New Zealand students, 55% were in the *often* ELA category, which was the third highest percentage of all countries. However, there was a marked difference of 38 scale points in mean reading achievement between New Zealand students in the *often* and *sometimes* categories, a difference that was larger than all but four countries (Trinidad/Tobago, Malta, United Arab Emirates, and Oman) and much larger than the international mean difference of 23 points (there were insufficient numbers of students in the lowest categories of the ELA and PLR scales to make reliable comparisons).

The second measure of literate capital used in the PIRLS 2011 study was the PLR scale, which was based on parents' degree of agreement/disagreement with seven statements regarding their attitude toward reading (e.g., I read only if I have to, reading is an important activity in my home, I like to spend my spare time reading, I read only if I need information). Children were assigned to three categories on the PLR scale (*like*, *somewhat like*, *do not like*) according to the parent's combined responses to the seven statements. Internationally, the

mean percentage of students in each category was 32% for the *like* PLR category, 57% for the *somewhat like* category, and 11% for the *do not like* category. Among New Zealand students, 51% were in the *like* category, which was the 2nd highest percentage of all countries. However, there was a marked difference of 40 scale points in mean reading achievement between New Zealand students in the *like* and *somewhat like* categories, a difference that was exceeded by only two other countries (United Arab Emirates and Morocco) and was much larger than the international mean difference of 28 points.

Overall, the results from the two measures of literate cultural capital were very similar. The difference in future reading achievement between New Zealand students in the high and middle categories of each scale was much larger than that of most other countries, very similar to what was reported in examinations of the PIRLS 2001 and 2006 data (Tunmer et al., 2004; Tunmer et al., 2008). These highly consistent findings suggest that the constructivist, multiple cues approach to teaching reading in New Zealand is generally beneficial to children with large amounts of literate cultural capital at school entry, but much less so for children with more limited amounts, which would explain the relatively high level of disparity among New Zealand readers in later grades.

Consistent with these findings are data from the PIRLS 2011 study on the mean reading achievement scores of students attending schools grouped into three categories: *more affluent* (schools where more than 25% of students come from economically affluent homes and not more than 25% from economically disadvantaged homes), *neither more affluent nor more disadvantaged*, and *more disadvantaged* (schools where more than 25% of students come from economically disadvantaged homes and not more than 25% from economically affluent homes). The percentages of New Zealand students in the three categories of schools were 39%, 34%, and 27%, respectively, and the mean reading scores were 560, 533, and 489, indicating a positive relationship between socioeconomic factors (i.e., decile level of the school) and reading achievement. A similar pattern of results was observed in every country, and internationally the mean reading scores were 530, 515, and 490. As noted previously, children from more economically advantaged backgrounds typically possess more literate cultural capital at school entry than children from low-income backgrounds. Given this relationship, we expected to find a relatively large difference in reading achievement between New Zealand students attending schools with proportionately few economically disadvantaged student and those attending schools with proportionately many. There was a difference of 71 scale points in mean reading performance between schools in New Zealand of high and low socioeconomic composition, a difference that was exceeded by only one country (Columbia), and was much larger than the international mean difference of 40 points. This pattern of results was very similar to what was reported in the PIRLS 2001 and 2006 studies, indicating that the relatively large gap in reading achievement between high and low decile schools in New Zealand has not changed during the past decade.

Specified differences in literate cultural capital at the beginning of school were associated with larger differences in later reading achievement in New Zealand than in most other countries despite the fact New Zealand was generally well above the international average in the availability of school resources for teaching reading, school climate, level of formal education achieved by reading teachers, and classroom learning environment. Data reported in the PIRLS 2011 study indicated that instruction in New Zealand was not affected by reading resource shortages, that teacher working conditions were good, that schools were well supplied with books and computers, that there was a very high emphasis on academic success (including teachers' expectations for student achievement), that schools were

generally safe and orderly, that teacher career satisfaction was above average, that collaboration among teachers to improve teaching was strong, and that amount of instructional time spent on teaching reading was the fourth highest of participating countries. An issue that was not investigated in the PIRLS 2011 study was the relationship between reading achievement and the approach to teaching reading adopted by countries. This probably would have been a difficult undertaking as within country variation in approaches to literacy instruction is often high. As noted previously, the approach to literacy education in New Zealand is relatively uniform.

Restrictive policies regarding the first year of literacy teaching

Reading Recovery

Wilkinson et al. (2000) stated that for major changes to occur in New Zealand's approach to teaching reading, "New Zealand educators will need to resolve the tension between explicit instruction and a developmental constructivist bias" (p. 12). However, little or no progress has occurred in resolving this tension as there continues to be strong resistance to providing beginning readers with assessment and explicit instruction in skills that are essential for reading development (e.g., phonological awareness, alphabetic coding skills), especially during the first year of schooling when instruction in these skills would be most effective (Lonigan & Phillips, 2012; Wagner, 2008).

The first formal assessment of literacy skills in New Zealand occurs at the end of the child's first year of schooling with the use of the Observation Survey developed by Clay (1998). Clay (2005a) argued that this in-depth assessment (which is part of the information used to determine eligibility for RR) should not occur until the end of the child's first year of formal instruction because "the child should be given sufficient time to adjust to the school situation and a variety of opportunities to pay attention to literacy activities" (p. 12). However, in her earlier work Clay (1979) made the important observation that some beginning readers find it extraordinarily difficult to hear the component sounds of spoken words. For this reason Clay (1985) incorporated into her RR programme a phonological awareness training procedure based on the work of the Russian psychologist, Elkonin (1973), who used "sound boxes" to teach children to isolate individual sounds in familiar printed words.

But the question that can be asked is this: why wait an entire year before helping children with little or no phonological awareness at school entry? A more effective strategy for improving the developmental trajectories of these beginning readers would be to intervene at an earlier point (Lonigan & Phillips, 2012). Wagner (2008) argued against a "wait-to-fail" approach to reading intervention and recommended that new entrants should receive an initial evaluation consisting of measures of emergent literacy skills that are known to be important in early literacy development (e.g., phonological awareness, print awareness). Supplementary instruction in these skills would then be provided to those children who needed it.

Constrained Skills Theory

Included in the description of the theoretical basis of the recently introduced national literacy standards described previously is a highlighted quote from Paris (2005) and a discussion of the distinction between "constrained" and "unconstrained" skills (Ministry of Education, 2009, p.9). Constrained skills include phonological awareness, alphabetic coding skills, and reading fluency (i.e., automaticity in word recognition). Skilled readers ultimately master

these skills completely and in a relatively short period of time. In contrast, unconstrained skills develop more slowly and are never completely mastered, as they continue to develop over the course of a lifetime. Unconstrained skills include vocabulary and comprehension skills.

Constrained Skills Theory (CST; Paris, 2005; Paris & Luo, 2010) argues that the distributional properties of constrained skills render traditional parametric statistical analyses inappropriate, in which case much of the research on the role of these skills in early reading development "may need to be reinterpreted" (Paris, 2005, p. 187). According to CST, the fact that all readers ultimately master constrained skills indicates that these skills have only limited importance and are not causally related to the development of reading ability. They should therefore not be the focus of early literacy instruction.

Although not explicitly stated, the MoE appears to be using the distinction between constrained and unconstrained skills to avoid including in the national literacy standards an assessment of the core skills that underlie the development of literacy (Greaney & Tunmer, 2010). Instead, the literacy standards document focuses largely on describing sets of comprehension-enhancing strategies that students should be able to demonstrate at each year level. Teachers are not required or encouraged to undertake systematic assessments of constrained skills at any point during the primary school years. The Observation Survey (Clay, 1998) administered to children after they have completed a year of formal schooling does not include measures of phonological awareness, alphabetic coding skills, or reading fluency. The procedure used to assign text level in the Observation Survey is based on word recognition accuracy only.

Paris and Luo (2010) correctly argued that focusing attention on teaching constrained skills during the early stages of literacy acquisition should not come at the expense of emphases on the development of vocabulary and comprehension skills. Research indicates that vocabulary knowledge at the beginning of school not only appears to have an immediate impact on the development of word recognition skills but also has a strong direct relation to future reading comprehension performance (Sénéchal, Ouellette, & Rodney, 2006; Tunmer & Chapman, 2012a,b). Children with limited understanding of the words of spoken language will encounter difficulty constructing meaning from text. During the early stages of learning to read, oral language factors such as vocabulary knowledge do not "show up" as major influences on reading comprehension because the inability to recognize the words of text limits the ability to understand text. However, this does not suggest that instruction in unconstrained skills should be delayed until children have acquired fast, accurate word recognition skills (Tunmer & Chapman, 2012b).

The broader claims of CST, however, have been challenged (Lonigan & Phillips, 2012; Schatschneider & Lonigan, 2010). Schatschneider and Lonigan argued that CST should be rejected on conceptual, statistical, and empirical grounds as it is "based on a flawed understanding of statistics, an idiosyncratic notion of causation, and assertions that are not borne out by the evidence" (p.347). Regarding causation, a large body of scientific research indicates that comprehending text in an alphabetic orthography depends on the ability to recognize the words of text accurately and quickly; that the development of automaticity in word recognition in turn depends on the ability to make use of letter-sound relationships in identifying unfamiliar words; and that the ability to discover mappings between spelling patterns and sound patterns in turn depends on the ability to detect phonemic sequences in spoken words (Pressley, 2006). Paris and Luo (2010) agree that phonological awareness,

alphabetic coding skills, and automaticity in word recognition "are enabling and necessary skills" (p.317). But as Schatschneider and Lonigan pointed out, this conceptualization of constrained skills satisfies the basic definition of *cause* as "that which brings something else into being" (p.350).

An important issue overlooked by CST is that ongoing weaknesses in essential constrained skills during the early stages of literacy acquisition can be developmentally limiting (Stanovich, 1986). Students with poorly developed constrained skills during the early stages of learning to read may eventually overcome their weaknesses and gradually develop along normal lines as they grow older. However, as suggested by the development of negative Matthew effects described previously, a more likely possibility is that children who do not possess sufficient levels of constrained skills (and who are not provided with explicit instruction to develop these competencies) will not await the development of these skills but instead will rely increasingly on ineffective compensatory strategies to identify unfamiliar words in text (such as using picture cues, partial visual cues, and contextual guessing). The continued use of these strategies inevitably leads to (causes) literacy learning difficulties, avoidance of reading, inattentive behaviours, low expectations of success, and withdrawal from literacy learning tasks. These students may rely on ineffective strategies to such an extent and for such a long period of time (years in some cases, especially when the use of these strategies is strongly encouraged in classroom instruction) that the strategies become entrenched and very difficult to "unlearn", even though the students may have acquired the necessary constrained skills for developing more effective word identification strategies (e.g., letter knowledge, phonological awareness).

Regarding the MoE's policy decisions based on CST, perhaps the most important point the MoE needs to consider is what Paris (2005) actually said about constrained skills. He stated that beginning readers "need to be instructed on those skills *early and persistently* by teachers and parents. Constrained skills *must be mastered*" (p.199, emphasis added). Paris further claimed that "constrained skills need to be mastered because they are necessary but not sufficient for other reading skills. They enable automatic decoding, deployment of attention, and application of comprehension strategies so they *set the stage for reading development*" (p. 200, emphasis added).

These comments strongly suggest that the MoE needs to place much greater emphasis on the assessment and teaching of constrained skills (i.e., phonological awareness, alphabetic coding skills, automaticity in word recognition) during the first year of formal schooling. However, policy seems to be moving in the opposite direction. The MoE recently announced a request for proposal (RFP) on *Continuity of early learning: Learning progress and outcomes in the early years*, with particular attention focused on the transition from early childhood education to school (Ministry of Education, 2013). For children aged 0-8 years, the RFP categorically states that "quantitative assessment is not a reliable method of collecting data for this age group" (p.4). This statement conflicts with a huge scientific literature on human development in general, and literacy development in particular.

Culturally responsive instruction

Another source of the strong resistance to early skills-based teaching is the assumption that this approach to reducing the large inequities in achievement outcomes in literacy education is based on "deficit theory" that "pathologizes" children from low-income, culturally diverse backgrounds and is antithetical to culturally responsive instruction (Berryman & Bishop,

2011; Fayden, 2005; Harris, 2009; McNaughton, 2004). Harris stated that words such as *gap, underachievement, disparity*, and *at-risk* "signal perceived deficiencies" (p.12). Berryman and Bishop argued that, "If we see students as having deficiencies, then our practices will address deficiencies" (pp. 250-251). In a critique of MoE policy and practice, Cullen (2007) pointed out a major anomaly in the general acceptance of this view by the MoE: "The Reading Recovery programme is essentially a withdrawal approach to remedy reading delays and deficits, yet at the same time classroom teachers are being progressively inducted into a co-constructivist perspective that emphasizes a credit perspective on family and community literacy meanings and practices" (p. 120). Clark (2006) suggested that the deficit connotations associated with skills-based teaching reflect the growing influence of cultural relativism in New Zealand education.

The views expressed by Berryman and Bishop (2011), Harris (2009), and others appear to question the widely held view that learning to read is a developmental process that takes place over time, involves qualitatively different (but perhaps overlapping) phases, and may break down at different points due to the failure to acquire the core skills that underlie the development of literacy (Ehri, 2005; Pressley, 2006; Snow & Juel, 2005; Tunmer & Nicholson, 2011). The literacy learning needs of children necessarily vary because they differ in the amount of reading-related knowledge, skills, and experiences (i.e., literate cultural capital) they bring to the classroom, in the explicitness and intensity of instruction they require to learn skills and strategies for identifying words and comprehending text, and in their location along the developmental progression from pre-reader to skilled reader. Prochnow, Tunmer, and Chapman (2012) recommended replacing each occurrence of the word *deficiencies* with the phrase *specific learning needs* in the quote from Berryman and Bishop (2011) in the preceding paragraph so that it reads: "If we see students as having *specific learning needs*, then our practices will address *specific learning needs"* (p. 215).

Regarding culturally responsive instruction, there seems to be little disagreement that teachers should adjust their teaching to accommodate student differences in cultural/family background, prior knowledge and experience, academic ability, personality, motivation, and preferred ways of doing things. Examples of culturally responsive instruction include making learning experiences more personally meaningful to students of diverse backgrounds by engaging them in activities that are related to their interests and experiences outside of school; using instructional materials that present minority cultures in an authentic manner, including presenting culturally relevant content in culturally familiar social contexts; taking into consideration culturally based learning styles; and improving community involvement in literacy learning by promoting stronger connections among schools, parents, and the community. Equally important, teachers should hold high expectations for students from culturally diverse backgrounds. Au (2000) argued that "there should not be a different set of standards for students of diverse backgrounds, but there should be recognition that these students may require more powerful instruction and additional time to meet the standards" (p. 844).

Reducing the gap in literacy achievement requires that teachers address the specific needs of children struggling to learn to read in an alphabetic orthography *regardless* of the cultural group or social class to which they belong. There is no evidence to suggest that explicit instruction in essential reading-related skills (i.e., constrained skills) cannot be done in a culturally sensitive and responsive manner. Teachers should certainly adjust their teaching to support students' identities, while at the same time holding the same expectations and

standards of achievement for students of diverse backgrounds as for those from the dominant culture.

Support for this view comes from one of the strongest proponents of culturally responsive instruction, Katherine Au (1998, 2000), who drew attention to the disadvantages of using the constructivist approach to teaching reading with students of diverse backgrounds. Au (1998) argued that a constructivist orientation "fails to acknowledge that a given set of learning opportunities may benefit mainstream students while working to the detriment of students of diverse backgrounds within the same classroom" (p. 307). She further stated that,

> Because the emphasis in constructivist approaches tends to be on process rather than product, educators with a mainstream constructivist orientation may see it as their role to act as facilitators of students' learning, responding to students' work but not transmitting knowledge … Educators with this orientation may be reluctant to provide students with instruction on specific skills. (p. 313)

Prochnow and Macfarlane (2011) expressed a similar view, stating that, "Culturally responsive teachers will recognize that learners with diverse learning needs begin school with limited experiences and limited knowledge of the fundamental building blocks that facilitate further learning, and these teachers will respond with instruction that targets the skill gaps and allows the learners to progress with their peers" (p. 162).

Summary

In summary, we have argued in this section that New Zealand's relatively large literacy achievement gap can be explained in terms of literate cultural capital and a constructivist orientation toward literacy education. The research we have discussed indicates that children enter school with large individual differences in the skills and competencies (i.e., literate cultural capital) important in learning to read; that children from low-income and/or culturally diverse backgrounds on average begin school with considerably lower levels of literate cultural capital than middle-class children; that children who possess higher levels of literate cultural capital at the beginning of school generally profit more from literacy instruction, learn to read sooner, and read better than those who do not; and that according to data from the PIRLS 2001, 2006, and 2011 studies, specified differences in literate cultural capital at the beginning of school are associated with larger differences in future reading achievement in New Zealand than in most other countries. Given these findings, the challenge for policy makers in the MoE is to develop an approach to literacy education in which the new entrant with limited literate cultural capital has approximately the same probability of success in learning to read and write as the new entrant with an abundance of literate cultural capital; that is, an approach that does not contribute to cultural reproduction in New Zealand society. Such an approach is described in the next section of the paper.

WHAT CAN BE DONE TO OVERCOME THE FAILURE OF NEW ZEALAND'S NATIONAL LITERACY STRATEGY?

Our recommended strategy for reducing the large inequities in achievement outcomes in New Zealand literacy education is to make fundamental changes to regular classroom literacy instruction and to replace RR with an alternative intervention programme that is specifically designed to target those struggling readers who need help the most.

Classroom literacy instruction

In view of the unacceptably large gap in literacy achievement in New Zealand, the critical question that needs to be addressed is how classroom literacy instruction can be changed to reduce the influence of differences in literate cultural capital at school entry on future reading achievement. We claim that the most effective strategy for reducing the literacy achievement gap is to use differentiated instruction from the outset of formal schooling that takes into account child-by-instruction interactions (Arrow & Tunmer, 2012; Tunmer & Nicholson, 2011).

According to contemporary learning theory, any act of learning is the joint product of the learner and the environment. In developing a conceptual framework for theories of learning to read, Byrne (2005) argued that different acts of learning can be located at different points along a continuum representing the division of labor between the learner and the environment. These ideas are represented graphically in Figure 1. The right end of the division of labor continuum represents acts of learning requiring only meager and fragmentary environmental input for learning to occur (such as learning spoken language), whereas the left end represents learning that requires rich and highly structured input from the environment (such as learning geometry).

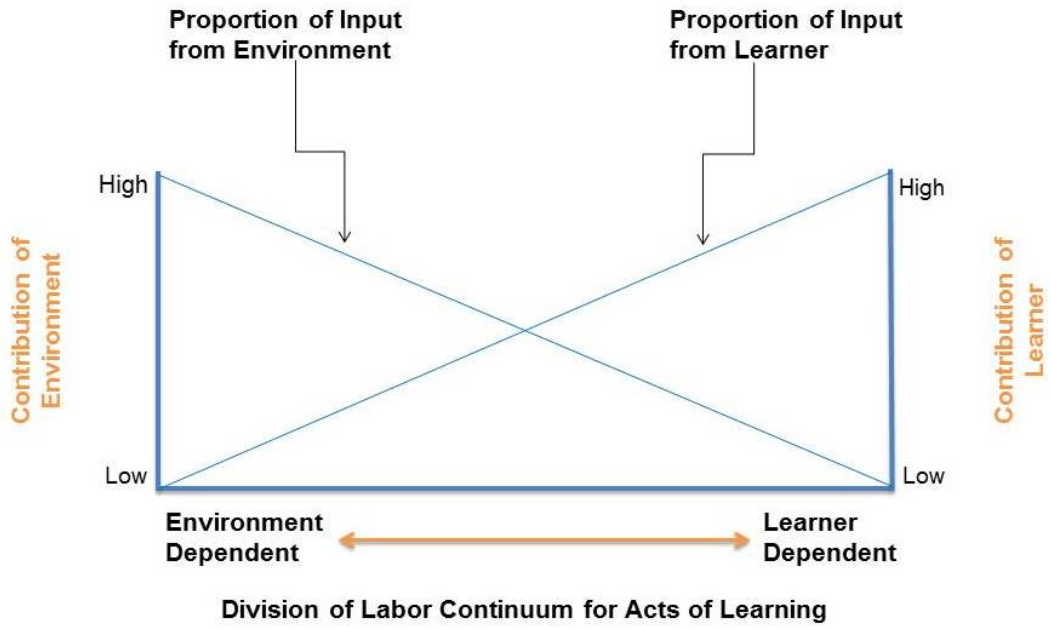

Figure 1. Theoretical continuum representing the division of labor between the learner and the environment.

As noted previously, children must rely increasingly on implicit learning to acquire the alphabetic coding skills necessary for learning to read. In terms of Byrne's (2005) division of labor continuum, the processes of learning to recognize words are initially *environment dependent* but necessarily become increasingly *learner dependent*. Research indicates, however, that the amount of explicit instruction in phonemic awareness and letter-sound relationships needed to initiate the process of inducing sublexical relations appears to vary considerably across children (Snow & Juel, 2005). For some beginning readers, the processes of acquiring literacy skills are highly *learner dependent.* They seem to grasp the idea of what

is required to discover orthographic patterns after having had only a small amount of phonologically-based skills and strategies explicitly taught to them. In contrast, for other children the learning processes are more *environment dependent*, with the children requiring a fairly structured and teacher-supported introduction to reading.

A key question is what causes these individual differences in literacy learning processes? Research suggests that the location of children on Byrne's (2005) division of labor continuum at school entry depends largely on the amount of literate cultural capital they possess, with *learner-dependent* children (typically from more advantaged backgrounds) having higher levels of essential reading-related knowledge, skills, and experiences, and *environment-dependent* children (typically from low-income backgrounds) having more limited amounts (see Figure 2). In a study examining the effects of different instructional emphases on children possessing varying amounts of literacy-related skills at the beginning of school, Juel and Minden-Cupp (2000) found that children with limited school-entry literacy skills benefitted more from explicit, code-emphasis approaches to beginning reading instruction than from whole language, "book experience" approaches, whereas the opposite pattern occurred with children who had higher levels of literate cultural capital at the beginning of school. Perhaps most importantly, Juel and Minden-Cupp reported that "the classroom… that had the very highest success both overall and with the low group had considerably different instruction across the groups" (p. 482).

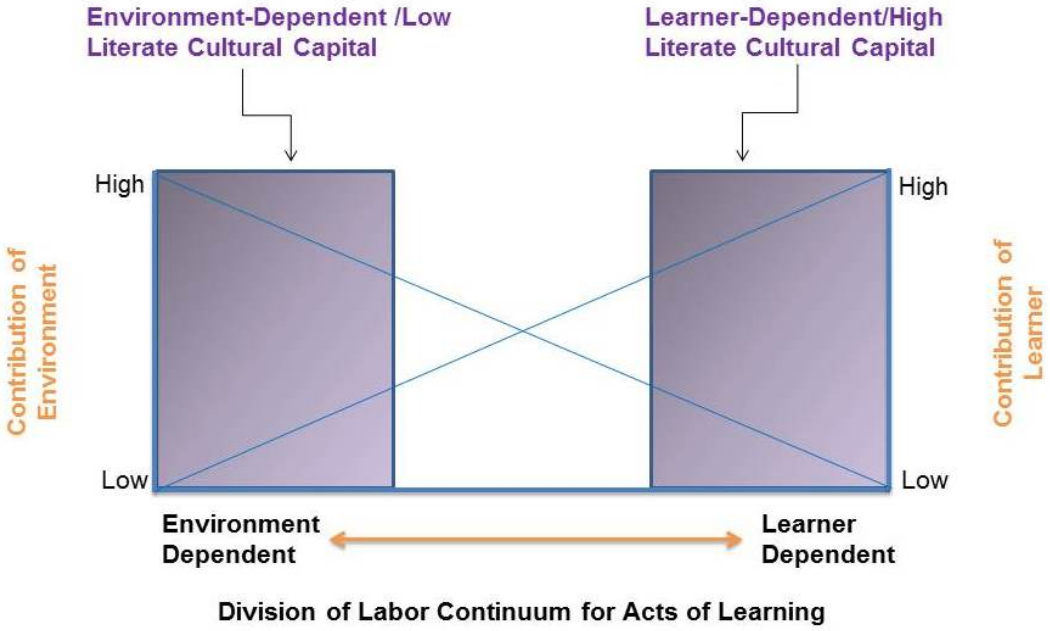

Figure 2. Learner-dependent children have higher levels of literate cultural capital at school entry, whereas environment-dependent children have more limited amounts.

The findings reported by Juel and Minden-Cupp (2000) suggest that a predominantly constructivist, book experience approach to reading instruction with a major emphasis on reading books, writing text, and some incidental teaching of word analysis skills during reading and writing activities is likely to be more suitable for *learner-dependent* children than heavy code-emphasis approaches. However, *environment-dependent* children will almost certainly benefit more from beginning reading instruction that includes explicit, systematic teaching of phonological awareness and alphabetic coding skills outside the context of

reading text in combination with plenty of opportunities to practice and receive feedback on using these skills during text reading. Providing *environment-dependent* beginning readers with explicit instruction in word analysis skills outside the context of reading text helps to ensure that these children see the importance of focusing on word-level cues as the most useful source of information in identifying words, as well as helping them to overcome any tendency they may have to rely mostly on sentence context cues to identify unfamiliar words in text rather than using context to supplement word-level information. Explicit instruction in word identification strategies does not require the implementation of a curriculum that is rigid, fixed, and lock-step, with the same lesson given to every child. As Lonigan and Phillips (2012) pointed out, "Virtually all of the evidence supporting explicit instruction has involved small-group, brief, hands-on activities that fit well with models of differentiated, scaffolded instruction…with warm, sensitive interactions, and with the appropriateness of varying group sizes and modalities" (p. 154).

Supporting an interaction between student characteristics (i.e., high vs. low literate cultural capital) and method of teaching reading (i.e., constructivist vs. explicit approaches) are the results of studies by Connor and colleagues (Connor, Morrison, & Katch, 2004; Connor et al., 2009). Connor et al. (2004) studied beginning readers who varied in the amount of literacy-related skills they possessed at school entry, which included letter identification, letter-sound correspondence, and oral vocabulary. In addition to print-related skills, vocabulary is important during the early stages of learning to read because children with poorly developed vocabulary knowledge will have trouble identifying and assigning appropriate meanings to unknown printed words, especially partially decoded or irregularly spelled words, if the corresponding spoken words are not in their listening vocabulary. This in turn will limit the development of their alphabetic coding skills, as additional spelling-sound relationships can be induced from words that have been correctly identified (Tunmer & Chapman, 2012a,b).

Conner et al. (2004) found that children who began first grade with below-average reading-related skills made larger reading gains in classrooms that provided greater amounts of teacher-managed, code-focused instruction throughout the year than in classrooms that provided greater amounts of child-managed, meaning-focused instruction. In contrast, for children with high levels of reading-related skills at school entry, greater growth in reading was achieved in classrooms that provided lesser amounts of teacher-managed, code-focused instruction and greater amounts of child-managed, meaning-focused instruction. Of particular importance was the finding that when student characteristics were appropriately matched with instructional approach, the improvement in end of year reading scores resulting from good fitting instructional patterns varied greatly between children with high and low levels of literate cultural capital at the beginning of first grade. For high literate cultural capital children, better fitting instructional patterns (i.e., child-managed, meaning-focused instruction) resulted in about half a grade equivalent gain in end of year reading scores over less effective instructional patterns. However, for low literate cultural capital children, better fitting instructional patterns (i.e., teacher-managed, code-focused instruction) resulted in a difference of more than *two full grade equivalents* in end of year reading scores compared with poorly fitting instructional patterns. These findings have particular significance for literacy education in New Zealand, as the relatively high level of disparity in reading achievement outcomes among New Zealand readers can be explained in terms of a largely unidimensional, constructivist approach to literacy teaching that produces poorly fitting instructional patterns for students with limited amounts of literate cultural capital at the beginning of school.

The results of the Connor et al. (2004) study indicate that instructional strategies that may be effective with some students may be less effective when applied to other students with different skills. In support of this claim, Conner et al. (2009) reported that children in first-grade classrooms that individualized reading instruction by taking into account child-by-instruction interactions made greater gains in reading achievement than children in control classrooms. The research reported by Connor and colleagues has important implications for the teaching of reading and preservice teacher training. To ensure that beginning reading instruction is optimally effective, greater emphasis needs to be placed on differentiated instruction, where teachers use research-based assessment procedures and instructional strategies to cater to the differing skill needs of beginning readers from the outset of schooling, with particular attention focused on ensuring the development of phonemically-based word-level skills and strategies by all children during the early stages of reading acquisition.

The ability to determine what instructional approach works best for which students requires high levels of teacher knowledge and professionalism. However, many practicing teachers exhibit weaknesses in concepts pertaining to the structure of language and nature of English orthography (Brady, 2011; Cunningham, Perry, Stanovich, & Stanovich, 2004; Moats & Foorman, 2003; Piasta, Connor McDonald, Fishman, & Morrison, 2009), including teachers, remedial specialists, and teacher trainees in New Zealand (Carroll, Gillon, & McNeill, 2012). Research also suggests that the instructors responsible for the training of preservice and in-service teachers are generally not well informed about the body of knowledge required to teach systematically phonological awareness and alphabetic coding skills (Binks-Cantrell, Washburn, Joshi, & Hougen, 2012; Joshi et al., 2009). Focusing greater attention on enhancing the quality of teacher preparation and professional development is likely to be an important aspect of increasing the effectiveness of reading instruction in New Zealand.

Reading Recovery

From our previous examination of data from the latest RR annual monitoring reports, we concluded that RR has had little or no impact on reducing New Zealand's relatively large literacy achievement gap. There are serious shortcomings and much needed improvements in several aspects of RR, including the theoretical underpinnings of the program, the assessment battery used in the program, the specific procedures and instructional strategies emphasized in the program, the manner of programme delivery (one-to-one instruction versus instruction in pairs), and congruence between classroom curriculum and the RR program. Fundamental changes in all of these areas would very likely improve the effectiveness of the program, both in terms of outcomes and cost (Church, 2005; Reynolds & Wheldall, 2007; Tunmer & Chapman, 2003, 2004).

The most serious shortcoming of RR, however, concerns the differential effectiveness of the program. As noted previously, the programme is beneficial for some struggling readers but not others, especially those struggling readers who need help the most. For these children, 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, Tunmer, & Chapman, 2005; Tunmer & Greaney, 2008, 2010). Given these considerations, the Literacy Experts Group (Ministry of Education, 1999a) that advised the Literacy Taskforce (Ministry of Education, 1999b) described previously included in its report 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" (p. 6).

Although the Literacy Taskforce did not adopt this recommendation, it did recommend a review of the RR programme (p. 23). However, as indicated in an article by Rivers (2001, February 16), who interviewed the developer of RR, making significant changes to RR based on the review would be difficult:

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

The RR programme is currently overseen by the Marie Clay Literacy Trust (http://irrto.us/index.php/marie-clay-literacy-trust), which is responsible for the copyright of all RR materials and the RR trademark. No changes in the materials or procedures of RR can therefore be made without the approval of the trustees. This makes it virtually impossible for school systems or countries to make changes to the RR programme based on recent research or to conduct independent studies investigating ways of modifying the programme to improve outcomes and/or cost effectiveness.

In a study of the effectiveness of RR, McDowall et al. (2005) found that RR was less beneficial to Māori and Pasifika students than to other students. Problems associated with the benefits of RR for Māori and Pasifika were generally attributed to implementation, resourcing, family/cultural factors, and inappropriate textual materials but not to the programme itself. McDowall et al. overlooked the fundamental problem with RR, which is that it is based on the multiple cues theory of reading, a model of reading that was rejected by the scientific community over three decades ago (e.g., Stanovich, 1980). Church (2005) made a similar point, stating that RR "was designed in the 1970s prior to most of the modern research into how children learn to read. Not surprisingly, therefore, it lacks a number of elements which have been found by research to be essential in teaching low achieving children how to read" (p. 13). As part of the effort to overcome the failure of New Zealand's national literacy strategy, RR needs to be replaced with an intervention programme that is based on contemporary theory and research on reading intervention and targets children who are most at risk of failing to learn to read.

CONCLUSIONS

Based on an analysis of data from the PIRLS 2001 study, Tunmer et al. (2004) predicted that unless fundamental changes were made to New Zealand's literacy strategy, the results of the next PIRLS study would show "no substantial reduction" in the unacceptably large reading achievement gap between good and poor readers (p. 141). Unfortunately, this prediction has turned out to be true not only for the PIRLS 2006 results, but for the PIRLS 2011 results as well. The aim of the current study was to provide an in depth analysis of the factors that have contributed to these unsatisfactory outcomes. Our central claim is that little or no progress has been made in reducing the literacy achievement gap because the constructivist/multiple cues model of reading adopted by the MoE as the theoretical basis for its approach to literacy teaching and intervention is fundamentally flawed. Consistent with this claim is a report from the Education Review Office (2009) indicating that, relative to its size, New Zealand spends a considerable amount on professional development for teachers ($200 million per year, over

half of which is allocated to professional development providers). The results of the PIRLS studies suggest that these expensive professional development initiatives have had no lasting impact on reading achievement, most likely because the initiatives were not based on fundamental changes to the general approach to teaching literacy adopted by the MoE.

We began the paper by providing a brief summary of the key reports, reviews, and government policy initiatives that have occurred in New Zealand over the past 15 years to underscore the levels of concern expressed by educators and policy makers regarding the persistently large inequities in literacy achievement outcomes. The remainder of the paper was then divided into three sections. In the first section we presented evidence in support of our claim that the MoE's literacy strategy has failed, drawing on data from the PIRLS 2011 study (Mullis et al., 2012) and the latest annual monitoring report of RR data in New Zealand (Lee, 2011). Our examination of the PIRLS 2011 results revealed that virtually no changes in educational outcomes have occurred. The mean reading achievement score for New Zealand in the PIRLS 2011 study was not significantly different from the PIRLS 2001 or 2006 studies, and the standard deviation for New Zealand's reading scores was almost unchanged from the PIRLS 2001 and 2006 studies and exceeded the values of most other countries, including those of the six English-speaking comparison countries. The large differences in reading achievement scores between Pākehā/European and Māori/Pasifika students have also not changed over the past decade. Our analyses of RR data from annual monitoring reports and other sources indicated that RR has had little or no impact on reducing New Zealand's relatively large literacy achievement gap because the programme is of limited benefit to those students who need help the most, especially Māori/Pasifika students and students from low-income backgrounds. We also discussed research indicating that positive maintenance effects for the majority of students successfully discontinued from RR are modest or non-existent.

In the second section we presented arguments and evidence in support of what we claim are the major factors responsible for the persistence of New Zealand's comparatively wide gap in literacy achievement, and for why the gap has not diminished over the past 15 years despite major efforts by the MoE to address the problem. Three interrelated factors were identified as contributing to the failure of New Zealand's literacy strategy: a rigidly constructivist orientation toward literacy education, the failure to respond adequately to differences in literate cultural capital at school entry, and restrictive policies regarding the first year of literacy teaching.

In the third section we reviewed research in support of what we maintain is the most effective strategy for reducing the literacy achievement gap: the use of differentiated instruction from the outset of formal schooling that takes into account interactions between school entry reading-related skills (high vs. low literate cultural capital) and method of teaching reading (constructivist vs. explicit approaches). We also argued that RR needs to be replaced with an intervention programme that is based on contemporary theory and research on reading intervention and targets those struggling readers who need help the most.

The arguments and evidence that we have presented in this paper should not be taken as suggesting that New Zealand teachers are responsible for the failure of the national literacy strategy. As noted previously, data from the PIRLS 2011 study indicated that they are well above the international average in level of formal education and availability of school resources for teaching reading. Rather, the failure is largely the result of the misguided policy decisions of the MoE. In an examination of the interplay between policy, research, and practice in New Zealand literacy education, Cullen (2007) noted that "the teaching of reading

has been characterized by direct links between key researchers, policy, and practices since the adoption of Clay's Reading Recovery programme in the 1980s, based on a whole language, constructivist view of reading" (p. 119). She further stated that the "selective use of key researchers to guide curriculum developments has resulted in idiosyncratic and restricted use of national and international research on literacy, in New Zealand's literacy policies and practices" (p. 124). We agree with Cullen and conclude from our investigation of the factors contributing to the failure of the national literacy strategy that the MoE needs to abandon its functionally discriminatory approach to literacy education, an approach that turns differences in literacy-related knowledge, skills, and experiences at school entry into disadvantages.

At the beginning of our paper we mentioned the inflated claims of others regarding the relative merits of New Zealand's approach to literacy education. In closing, we wish to respond further to these claims by drawing attention to the following quote attributed to John F. Kennedy: "The great enemy of the truth is very often not the lie – deliberate, contrived, and dishonest, but the myth – persistent, pervasive, and unrealistic."

REFERENCES

Arrow, A. W., & Tunmer, W. E. (2012). Contemporary reading acquisition theory: The conceptual basis for differentiated reading instruction. In S. Suggate & E. Reese (Eds.), *Contemporary debates in childhood education and development* (pp. 241-249). London: Routledge.

Au, K. H. (1998). Social constructivism and the school literacy learning of students of diverse backgrounds. *Journal of Literacy Research, 20*, 297-319.

Au, K. H. (2000). A multicultural perspective on policies for improving literacy achievement: Equity and excellence. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook or reading research* (Vol. 3, pp. 835-851). Mahwah, NJ: Lawrence Erlbaum Associates.

Berryman, M., & Bishop, R. (2011). Societal and cultural perspectives through a Te Kotahitanga lens. In C. Rubbie-Davies (Ed.), *Educational psychology: Concepts, research and challenges* (pp. 249-267). London, UK: Routledge.

Binks-Cantrell, E., Washburn, E. K., Joshi, R. M., & Hougen, M. (2012). Peter effect in the preparation of reading teachers. *Scientific Studies of Reading, 16*, 526-536.

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.

Blaiklock, K. E. (2004). A critique of running records of children's oral reading. *New Zealand Journal of Educational Studies, 39,* 241-253.

Brady, S. (2011). Efficacy of phonics teaching for reading outcomes: Indications from post-NRP research. In S. Brady, D. Braze, & C. Fowler (Eds.), *Explaining individual differences in reading: Theory and evidence* (pp. 69-96). New York, NY: Psychology Press.

Byrne, B. (2005). Theories of learning to read. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 104-119). Oxford, UK: Blackwell.

Carroll, J., Gillon, G., & McNeill, B. (2012). Explicit phonological knowledge of educational professionals. *Asia Pacific Journal of Speech, Language and Hearing, 15,* 231-244.

Center, Y., Wheldhall, K., Freeman, L., Outred, L., & McNaught, M. (1995). An evaluation of Reading Recovery. *Reading Research Quarterly, 30,* 240-263.

Center on International Education Benchmarking (2012). *New Zealand*. Washington, DC: National Center on Education and the Economy. Retrieved from http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/new-zealand-overview/

Chamberlain, M. (2012). New Zealand. In I.V.S. Mullis, M. O. Martin, C. A. Minnich, K. T. Drucker, & M. A. Ragan (Eds.), *PIRLS 2011 encyclopedia: Education policy and curriculum in reading* (pp. 435-459). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Chamberlain, M., & Caygill, R. (2012). *Key findings from New Zealand's participation in the Progress in International Reading Literacy Study (PIRLS) and Trends in International Mathematics and Sciences Study (TIMSS) in 2010/11.* Wellington, New Zealand: Ministry of Education. Retrieved from http://www.educationcounts.govt.nz/publications/series/2539/114981/114991

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,* 17-29.

Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. (2001). Does success in the Reading Recovery programme depend on developing proficiency in phonological processing skills? A longitudinal study in a whole language instructional context. *Scientific Studies*

of Reading, 5, 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.
Clark, J. A. (2006). The gap between the highest and lowest school achievers: Philosophical arguments for downplaying teacher expectation theory. *New Zealand Journal of Educational Studies, 41,* 367-382.
Clay, M. M. (1979). *The early detection of reading difficulties*. Auckland, New Zealand: Heinemann.
Clay, M. M. (1985). *The early detection of reading difficulties*. Auckland, New Zealand: Heinemann.
Clay, M. M. (1987). Learning to be learning disabled. *New Zealand Journal of Educational Studies, 22,* 155-173.
Clay, M. M. (1991). *Becoming literate: The construction of inner control.* Auckland, New Zealand: Heinemann.
Clay, M. M. (1998). *An observation survey of early literacy achievement*. Auckland, New Zealand: Heinemann.
Clay, M. M. (2000). *Running records for classroom teachers.* Auckland, New Zealand: Heinemann.
Clay, M. M. (2002). *An observation survey of early literacy achievement* (2nd ed.). 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.
Connelly, V., Johnston, R., & Thompson, G. B. (2001). The effects of phonics instruction on the reading comprehension of beginning readers. *Reading and Writing, 14,* 423-457.
Connor, C. M., Morrison, F. J., & Katch, L. E. (2004). Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. *Scientific Studies of Reading, 8,* 305-336.
Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., Underwood, P., & Morrison, F. (2009). Individualizing student instruction precisely: Effects of child x instruction interactions on first graders' literacy development. *Child Development*, *80*, 77-100.
Crooks, T., & Caygi1l, R. (1999, December). *New Zealand's National Educational Monitoring Project: Māori student achievement, 1995-1998.* Paper presented at the combined annual conference of the New Zealand Association for Research in Education and the Australian Association for Research in Education, Melbourne, Australia.
Cullen, J. (2007). Literacy debate in the early years: The New Zealand context. In R. Openshaw & J. Soler (Eds.), *Reading across international boundaries: History, policy, and politics* (pp. 111-128). Charlotte, NC: Information Age Publishing.
Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Stanovich, P. J. (2004). Disciplinary knowledge of K-3 teachers and their knowledge calibration in the domain of early literacy. *Annals of Dyslexia, 54,* 139-167.
Department of Education (2010). *The new conceptual framework for teaching reading: The "simple view of reading"*. Retrieved from http://nationalstrategies.standards.dcsf.gov.uk/node/20162
Education Review Office (2009). *Managing professional learning and development in primary schools.* Wellington, New Zealand: Education Evaluation Reports. Retrieved

from http://www.ero.govt.nz/National-Reports/Managing-Professional-Learning-and-Development-in-Primary-Schools-May-2009
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* (pp. 135-154). Oxford, UK: Blackwell.
Elbaum, B., Vaughn, S., Hughes, M., & Moody, S. (2000). How effective are one-to-one tutoring programmes in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92*, 605-619.
Elkonin, D. R. (1973). USSR. In J. Downing (Ed.), *Comparative reading* (pp. 551-580). New York: Macmillan.
Elley, W. B. (1992). *How in the world do students read?* Hamburg, Germany: International Association for the Evaluation of Educational Achievement.
Elley, W. B. (2001). *STAR Supplementary Test of Achievement in Reading*. Wellington, New Zealand: New Zealand Council for Educational Research.
Fayden, T. (2005). *How children learn: Getting beyond the deficit myth*. Boulder, CO: Paradigm.
Flockton, L., & Crooks, T. (1997). *Reading and speaking assessment results 1996: National Education Monitoring Report 6.* Wellington, New Zealand: Ministry of Education.
Gilmore, A.M. (1998). *School entry assessment*. Wellington, New Zealand: Ministry of Education.
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.
Goldenberg, C. (2001). Making schools work for low-income families in the 21st Century. In S. B. Neuman & D. K. Dickinson (Eds.). *Handbook of early literacy research* (pp. 211-231). New York: Guilford Press.
Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6,* 126-135.
Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7,* 6-10.
Greaney, K. T. (2001). An investigation of teacher preferences for word identification strategies. *Australian Journal of Language and Literacy, 24*, 21-30.
Greaney, K. T., & Tunmer, W. E. (2010). The Literacy Learning Progressions and the Reading and Writing standards: Some critical issues. *Kairaranga, 11,* 23-27.
Harris, F. (2009). Reading-related language abilities: Māori children "at promise". *Set: Research Information for Teachers, 3*, 11-18.
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Brookes.
Hattie, J. A. (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*, 15-25.
Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2,* 127-160.
Iversen, S. A., & Tunmer, W. E. (1993). Phonological processing skill and the Reading Recovery program. *Journal of Educational Psychology, 85,* 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,* 456-472.
Joshi, R. M., Binks, E., Hougen, M., Dahlgren, M., Ocker-Dean, E., & Smith, D. (2009). Why elementary teachers might be inadequately prepared to teach reading. *Journal of Learning Disabilities, 42,* 392-402.
Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly, 35*, 458-492.
Korat, O., Klein, P., & Segal-Drori, O. (2007). Maternal mediation in book reading, home literacy environment, and children's emergent literacy: A comparison between two social groups. *Reading and Writing, 20,* 361-398.
Lee, M. (2011). *Annual monitoring of Reading Recovery: The data for 2011.* Wellington, New Zealand: Ministry of Education. Retrieved from http://www.educationcounts.govt.nz/publications/series/1547/111098
Lee, M. (2012). *Resource Teachers of Literacy Annual Report 2011.* Wellington, New Zealand: Ministry of Education. Retrieved from http://www.educationcounts.govt.nz/publications/series/2247/resource-teachers-literacy-annual-report-2011
Liberman, I. Y., & Liberman, A. M. (1992). Whole language versus code emphasis: Underlying assumptions and their implications for reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 343-366). Hillsdale, NJ: Erlbaum.
Limbrick, E. A., & Jesson, R. N. (2010, December). *Are gains from early literacy interventions sustained over subsequent years of primary schooling: The case of Reading Recovery in New Zealand.* Paper presented at the annual conference of the New Zealand Association for Research in Education, Auckland, New Zealand.
Lonigan, C. J., & Phillips, B. M. (2012). Understanding the contributions of early academic skills to children's success in school. In S. Suggate & E. Reese (Eds.), *Contemporary debates in childhood education and development* (pp. 145-158). London: Routledge.
McDowall, S., Boyd, S., & Hodgen, E., with van Vliet, T. (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.
McNaughton, S. (2004). Search for pearls of wisdom: A developmental perspective on the PIRLS results. *Delta, 56,* 21-31.
Ministry of Education. (1991). *Reading in junior classes*. Wellington, New Zealand: Learning Media Ltd.
Ministry of Education. (1996). *The learner as reader*. Wellington, New Zealand: Learning Media Ltd.
Ministry of Education. (1997a). *Adult literacy in New Zealand: Results from the International Adult Literacy Survey.* Wellington, New Zealand: Author.
Ministry of Education. (1997b). *Reading and beyond*. Wellington, New Zealand: Learning Media Ltd.
Ministry of Education. (1999a). *Literacy Experts Group: Report to the Secretary for Education.* Wellington, New Zealand: Author.
Ministry of Education. (1999b). *Report of the Literacy Taskforce*. Wellington, New Zealand: Author.
Ministry of Education. (2003a). *Effective literacy practice in Years 1 to 4.* Wellington, New Zealand: Learning Media Ltd.
Ministry of Education. (2003b, June 2). Learning to read. *New Zealand Education Gazette*, *82* (10), 8-10.

Ministry of Education. (2003c). *Sound sense: Phonics and phonological awareness.* Wellington, New Zealand: Learning Media Ltd.
Ministry of Education. (2009). *Reading and writing standards for Years 1-8*. Wellington, New Zealand: Learning Media Ltd.
Ministry of Education. (2010). *The literacy learning progressions.* Wellington, New Zealand: Learning Media Ltd.
Ministry of Education. (2011). *Briefing to the incoming Minister.* Wellington, New Zealand. Author.
Ministry of Education. (2013). Request for proposals for *Continuity of early learning: Learning progress and outcomes in the early years, two research projects.* Retrieved from http://www.gets.govt.nz
Moats, L. C., & Foorman, B. R. (2003). Measuring teachers' content knowledge of language and reading. *Annals of Dyslexia, 53,* 23-45.
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., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 International Report.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
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.
Nash, R. (1993). *Succeeding generations: Family resources and access to education in New Zealand.* Auckland: Oxford University Press.
Nash, R. (1997). Deficit theory and the family resource framework: Parkyn revisited. *New Zealand Journal of Educational Studies, 32,* 13-23.
National Reading Panel (2000). *Teaching children to read. An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute for Child Health and Human Development.
New Zealand Council for Educational Research. (2008). *PAT: Reading comprehension and vocabulary.* Wellington, New Zealand: Author.
New Zealand House of Representatives. (2001). *Report of the Education and Science Committee on the inquiry into the teaching of reading in New Zealand.* Wellington, New Zealand: Author.
New Zealand House of Representatives. (2008). *Inquiry into making the schooling system work for every child: Report of the Education and Science Committee.* Wellington, New Zealand: Author.
Nicholson, T. (1995). Research note: More news on rich and poor schools, and the news is still not good. *New Zealand Journal of Educational Studies, 30,* 227-228.
Nicholson, T. (1997). Closing the gap on reading failure: Social background, phonemic awareness, and learning to read. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 381-407). Mahwah, NJ: Lawrence Erlbaum Associates.
Nicholson, T. (1999). Literacy in the family and society. In G. B. Thompson & T. Nicholson (Eds.), *Learning to read: Beyond phonics and whole language* (pp. 1-22). New York: Teachers College Press.
Nicholson, T. (2003). Risk factors in learning to read. In B. Foorman (Ed*), Preventing and remediating reading difficulties: Bringing science to scale* (pp. 165-193). Timonium, MD: York Press.

Nicholson, T., & Gallienne, G. (1995). Struggletown meets Middletown: A survey of reading achievement levels among 13-year-old pupils in two contrasting socio-economic areas. *New Zealand Journal of Educational Studies, 30,* 15-24.

Paris, S. G. (2005). Re-interpreting the development of reading skills. *Reading Research Quarterly, 40,* 184-202.

Paris, S. G., & Luo, S. W. (2010). Confounded statistical analyses hinder interpretation of the NELP report. *Educational Researcher, 39*, 316-322.

Piasta, S. B., Connor McDonald, C., Fishman, B. J., & Morrison, F. J. (2009). Teachers' knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading, 13*, 224-248.

Pinker, S. (1994). *The language instinct: How the mind creates language.* New York: Harper Perennial.

Phillips, B. M., & Lonigan, C. J. (2005). Social correlates of emergent literacy. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 173-187). Oxford, UK: Blackwell.

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

Prochnow, J. E., & Macfarlane, A. H. (2011). Managing classroom behaviour: Assertiveness and warmth. In C. Rubie-Davies (Ed.), *Educational psychology: Concepts, research and challenges* (pp. 150-166). London, UK: Routledge.

Prochnow, J. E., Tunmer, W. E., & Chapman, J. W. (2012). Family resource theory, literate cultural capital, and literacy achievement. In R. Openshaw & J. Clark (Eds.), *Critic and conscience: Essays on education in memory of John Codd and Roy Nash* (pp. 203-221). Wellington, New Zealand: New Zealand Council for Educational Research.

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

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

Rose, J. (2006). *Independent review of the teaching of early reading: Final report.* London: Department of Education and Skills.

Ryder, J. F., Tunmer, W. E., & Greaney, K. T. (2008). Explicit instruction in phonemic awareness and phonemically-based decoding skills as an intervention strategy for struggling readers in whole language classrooms. *Reading and Writing, 21,* 349-369.

Scarborough, H. S., & Brady, S. A. (2002). Toward a common terminology for talking about speech and reading: A glossary of the *"Phon"* words and some related terms. *Journal of Literacy Research, 34,* 299-334.

Schatschneider, C., & Lonigan, C. J. (2010). Misunderstood statistical assumptions undermine criticism of the National Early Literacy Panel's Report. *Educational Researcher, 39*, 347-351.

Shankweiler, D., & Fowler, A. E. (2004). Questions people ask about the role of phonological processes in learning to read. *Reading and Writing, 17,* 483-515.

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

Smith, J. W. A., & Elley, W. B. (1997). *How children learn to read*. Auckland, New Zealand: Longman.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

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* (pp. 501-520). Oxford: Blackwell.

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

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 340-406.

Torgesen, J. K. (2004). Lessons learned from research on interventions for students who have difficulty learning to read. In R. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 355-382). Baltimore, MD: Brookes.

Tracey, D. H., & Morrow, L. M. (2006). *Lenses on reading: An introduction to theories and models.* New York: Guilford Press.

Tunmer, W. E., & Arrow, A. W. (2013). Reading: Phonics instruction. In J. Hattie & E. Anderman (Eds.), *International guide to student achievement* (pp. 316-319). London: Routledge.

Tunmer, W. E., & Chapman, J. W. (2002). The relation of beginning readers' reported word identification strategies to reading achievement, reading-related skills, and academic self-perceptions. *Reading and Writing, 15,* 341-358.

Tunmer, W. E., & Chapman, J. W. (2003). The Reading Recovery approach to preventive early intervention: As good as it gets? *Reading Psychology, 24,* 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* (pp. 99-114). Baltimore, MD: International Dyslexia Association.

Tunmer, W. E., & Chapman, J. W. (2012a). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading, 16*, 122-140.

Tunmer, W. E., & Chapman, J. W. (2012b). The Simple View of Reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities, 45,* 453-466.

Tunmer, W. E., Chapman, J. W., & Prochnow, J. E. (2003). Preventing negative Matthew effects in at-risk readers: A retrospective study. In B. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 121-163). Timonium, MD: York Press.

Tunmer, W. E., Chapman, J. W., & Prochnow, J. E. (2004). Why the reading achievement gap in New Zealand won't go away: Evidence from the PIRLS 2001 international study of reading achievement. *New Zealand Journal of Educational Studies, 39,* 127-145.

Tunmer, W. E., Chapman, J. W., & Prochnow, J. E. (2006). Literate cultural capital at school entry predicts later reading achievement: A seven year longitudinal study. *New Zealand Journal of Educational Studies, 41,* 183-204.

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* (pp. 241-267). London: Sage.

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

Tunmer, W. E., & Hoover, W. (1993). Components of variance models of language-related factors in reading disability: A conceptual overview. In M. Joshi & C. K. Leong (Eds*.), Reading disabilities: Diagnosis and component processes* (pp. 135-173). Dordrecht, The Netherlands: Kluwer Academic Publishers.

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, pp. 405-431). New York: Routledge.

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,* 105-119.

Tunmer, W. E., & Prochnow, J. E. (2009). Cultural relativism and literacy education: Explicit teaching based on specific learning needs is not deficit theory. In R. Openshaw & E. Rata (Eds.), *The politics of conformity in New Zealand* (pp. 154-190). Auckland, New Zealand: Pearson.

Tunmer, W. E., Prochnow, J. E., Greaney, K. T., & Chapman, J. W. (2007). What's wrong with New Zealand's national literacy strategy? In R. Openshaw & J. Soler (Eds.), *Reading across international boundaries: History, policy and politics* (pp. 19-42). Charlotte, NC: Information Age Publishing.

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

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,* 2-40.

Vellutino, F. R., Tunmer, W. E., Jaccard, J., & Chen, S. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11,* 3-32.

Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography.* New York, NY: Guilford Press.

Wagemaker, H. (Ed.). (1993). *Achievement in reading literacy: New Zealand's performance in a national and international context.* Wellington, New Zealand: Ministry of Education.

Wagner, R. K. (2008). Rediscovering dyslexia: New approaches for identification and classification. In G. Reid, A. J. Fawcett, F. Manis, & L. S. Siegel (Eds.), *The Sage handbook of dyslexia* (pp. 174-191). London: Sage.

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B. Neuman & D. K Dickinson (Eds.), *Handbook of early literacy research* (pp. 11-29). New York: Guilford Press.

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, pp. 3-16). Mahwah, NJ: Erlbaum.