MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

EARLY LITERACY RESEARCH PROJECT

Final Report to the New Zealand Ministry of Education

James W. Chapman, Alison W. Arrow, Christine Braid, Keith T. Greaney, and William E. Tunmer

College of Humanities and Social Sciences

Massey University

February, 2018

THE ENGINE OF THE NEW NEW ZEALAND

ABSTRACT

The main purpose of this longitudinal research project was to provide professional learning and development (PLD) workshops and associated materials for teachers of New Entrant/Year 1 students, and to assess the impact of the workshops on the literacy learning outcomes of students in Year 1 and beyond. The workshops focussed on providing teachers with the knowledge and skills to adopt explicit and systematic word-decoding teaching strategies in their literacy instruction. Effective word decoding skills are a necessary requirement for success in learning to read.

The project commenced in February 2015 and concluded in July 2017. A total of 729 students from 39 schools in the lower North Island participated in the project. Schools were randomly allocated to either an "intervention" group or a "comparison" group. Attrition, the withdrawal of one school, and incomplete data reduced the number of students included in the various analyses. Teachers who had some involvement with students in the project, either as a participating project teacher, a comparison school teacher, or as a classroom teacher in following years, numbered 288.

Two cohorts of teachers and students took part in the study. Cohort 1 students commenced school in February 2015. Literacy-related assessments were collected over 2½ years from school entry to the middle of 2017, when the students were in Year 3. Cohort 1 teachers took part in the PLD workshops during 2015. Cohort 2 students started school in February 2016. Literacy assessments were collected from school entry to the middle of 2017, when the students were in Year 2. Cohort 2 teachers participated in PLD workshops during 2016. Each cohort had a number of comparison schools, with teachers who continued with their normal literacy instruction. Students in these schools were also assessed.

Results for Cohort 1 students showed that the intervention was unsuccessful. Intervention students did not outperform Comparison students on any of the literacy assessments over the 2½ years of the project. Comparison students generally performed better than the Intervention students. Further, teachers showed no evidence of significant change in their knowledge of the language foundations associated with literacy teaching

and learning, or of change in literacy teaching practices. These findings were attributed to unexpected challenges teachers faced in implementing different teaching strategies.

Analyses of literacy assessments in the latter part of 2015 showing that Intervention students were not improving relative to Comparison students led to modifications to the PLD delivery and supporting materials. A second cohort of teachers and students was recruited from the 2015 Comparison schools that remained in the project.

Results for Cohort 2 Intervention students showed significantly better literacy learning outcomes than the Cohort 2 Comparison students at the end of Year 1 (2016), and in the middle of Year 2 (2017). Cohort 2 Intervention students also significantly outperformed Cohort 1 Intervention students at the mid-Year 2 assessment point on measures of reading and spelling. Especially significant was the finding that low decile Intervention students dramatically outperformed low decile Comparison students, and in some cases had mean literacy assessment scores that were close to or equal to those of students in higher decile schools.

Results for Cohort 2 Intervention teachers showed important improvements in their knowledge of the language foundations associated with effective literacy teaching and learning. Video clips of classroom teaching also revealed changes in instructional practices that reflected content and materials from the PLD workshops.

The significantly improved literacy learning outcomes for Cohort 2 Intervention students are consistent with previous research and with the theoretical basis of the project. Initial word learning, which forms the basis of reading comprehension, is dependent on students quickly mastering the ability to decode words in text. Many students benefit from systematic and explicit instruction in word-level decoding and in the language foundations associated with the alphabetic principle. Cohort 2 Intervention students clearly benefitted from changes their New Entrant/Year 1 teachers made to their literacy instruction as a result of the PLD workshops.

The results of this project show that there is a large potential benefit for teachers and students throughout New Zealand in changing significant aspects of literacy instruction in New Entrant and Year 1 classrooms.

ACKNOWLEDGEMENTS

The research presented in this report is the result of a team effort. In addition to the two Co-Principal Investigators, members of the team included Distinguished Professor Bill Tunmer, Dr Keith Greaney, Dr Jane Prochnow, and Christine Braid. Team members individually and collectively made significant contributions to the conceptual and theoretical basis of the project, the research design, the framework and content of the PLD workshops, and to many other factors associated with classroom instructional and assessment practices.

However, undertaking the project was not possible without the support of many other people. We are grateful to the teachers, principals, parents and students who participated in this research project.

We also acknowledge the contribution made by the research assistants. Their roles were complicated and involved learning how to administer the various assessments, how to negotiate access to schools, work with students, and often travel to locations some considerable distance from their homes.

We are particularly grateful for the significant contribution made by the Project Administrative Manager, Anne Palmer. Anne proved to be an efficient, diligent, conscientious, and highly competent assistant. She effectively managed the Project Office, project materials, liaison with schools and with Research Assistants, and management of data in preparation for analyses. Her unflappable manner, and extraordinary attention to detail was invaluable.

Finally, we are grateful to the Ministry of Education for its support of this project. Various Ministry officials have been involved at different stages from when the concept of the project was first considered in 2014. We appreciate the work done by those officials in supporting the idea of the project, and its progress during the nearly three years of data collection.

James Chapman and Alison Arrow
Co-Principal Investigators
Massey University
5 February, 2018

EXECUTIVE SUMMARY

This research project involved a two-cohort longitudinal study. The main purpose was to provide professional learning and development (PLD) workshops and associated materials for teachers of New Entrant/Year 1 students, and to assess the impact of the workshops on students' literacy performance in Year 1 and beyond. These workshops focussed on providing teachers with the knowledge and skills to adopt explicit and systematic word-level decoding teaching strategies in their literacy instruction.

Participants

A total of 729 students from 39 schools in the lower North Island participated in the project. Two hundred and eighty-eight teachers had some involvement during at least one of the 2½ years of the project, as either a PLD workshop participant or classroom teacher of participating students.

The project involved two cohorts of teachers and students. In 2015, 45 Cohort 1 teachers from 24 schools took part in five PLD workshops. Assessment data were collected on Cohort 1 students over a 2½-year period from close to school entry in February 2015 through to the middle of 2017, when the students were in Year 3. In 2016, 34 Cohort 2 teachers from 13 schools took part in the PLD workshops. Assessment data were collected from Cohort 2 students from close to school entry in February 2016 through to the middle of 2017, when the students were in Year 2.

Assessment data were also collected from students in schools whose teachers did not participate in the PLD workshops but which were willing to participate as points of comparison.

Teachers in both cohorts also took part in a range of literacy teaching-related survey and knowledge assessments. More extensive data were collected and analysed for Intervention teachers in Cohort 2.

Cohort 1 Students' Results

School entry assessment data revealed similarities between Intervention and Comparison students on most variables. Although not statistically significant, a number of

assessments showed Comparison students tended to obtain higher scores than Intervention students on key variables highly predictive of subsequent reading comprehension ability, such as letter name and letter sound knowledge.

Towards the end of Year 1, the Intervention students were not outperforming the Comparison students. One important contributing factor was the challenges teachers faced in changing their instructional practices in line with materials and demonstrations presented in the PLD workshops during Year 1. Another significant factor was the impact of teacher changes on students' literacy performances.

Of the 24 Intervention schools, seven made teacher changes during Year 1 that resulted in 45 Intervention students having a teacher who was not participating in the project PLD workshops. These changes affected students' literacy achievements.

All eight end of Year 1 assessments revealed highly statistically significant differences between Intervention students who had a project teacher throughout the year, compared to Intervention students who had a change to a non-project teacher during the year.

Skills emphasised in the PLD workshops (e.g., phonological processing, word knowledge) were markedly higher for "project teacher" students than "non-project teacher" students. Subsequent analyses of Intervention student assessment data excluded the students who were moved to a class taught by a non-project teacher.

Towards the end of Year 2, Comparison students obtained statistically significant higher scores than the Intervention students on measures of literacy foundations, word knowledge, and language comprehension.

At the end of the project in the middle of Year 3, the Comparison students continued to perform at higher levels than the Intervention students, but differences in scores were not statistically significant.

Cohort 1 Teachers' Results

Cohort 1 teachers' knowledge of basic language constructs associated with literacy teaching and learning was assessed on two occasions during 2015. We also assessed teachers' perceptions of their literacy teaching ability, their literacy-related teaching efficacy, and the types of prompts teachers use when students make word mistakes in reading.

Fifty-five Intervention and Comparison teachers participated in the first survey. No statistically significant differences were observed between Intervention and Comparison teachers. Many questions, however, were not answered, and in a number of cases, whole sections were incomplete.

Only 24 teachers participated in the end of Year 1 survey. Results revealed that Intervention teachers' knowledge of the language foundations associated with literacy teaching and learning had not improved relative to the Comparison teachers.

Teacher video data for Intervention teachers engaging in small group instruction showed virtually none of the changes to literacy teaching that would be expected as a result of the PLD workshops. Although there was some teaching of word knowledge, no teacher was seen to teach students to apply this knowledge in the reading process.

The observations are consistent with the lack of positive improvement in teacher knowledge. More importantly, the video data revealed reasons as to why Intervention students did not show improvements in literacy learning outcomes when compared to Comparison students. Teacher knowledge and instructional practices appeared to remain largely unchanged during Year 1.

Cohort 2 PLD Changes

As a result of the unexpected results for Intervention students and teachers during the first year of the project, we made changes to the materials and delivery of the PLD workshops, and recruited a second cohort of teachers and students.

The re-developed PLD programme provided more systematic guidance on how to implement word-level instructional strategies into effective practice. The workshops were changed to provide much more explicit implementation of the very specific teaching strategies designed to support the development of students' word-level skills in reading.

An explicit "how-to" guide was prepared that provided a scope-and-sequence framework within which teachers could identify the learning needs of their students in one of four developmental phases.

In addition, a teachers' "coach" was engaged to undertake observations of teachers providing literacy instruction, and to offer advice designed to further improve implementation of strategies from the PLD workshops. The coach visited almost all Cohort 2 teachers on four occasions.

Cohort 2 Participants

Thirty-eight schools from the 2015 sample agreed to continue with the project during 2016. Eleven schools with 27 teachers and 116 students formed the Cohort 2 Intervention group. Teachers in this group had been in Cohort 1 Comparison schools in 2015.

An additional Intervention group was formed: Intervention^{+}. In this group were 40 teachers from 20 schools who participated in the Cohort 1 PLD workshops; 135 students were in these teachers' classrooms in 2016. We reasoned that the teachers in this group may have more fully "digested" the material from the workshops, leading to improved implementation of the new strategies.

Eight teachers with 59 students from five schools agreed to form a Cohort 2 Comparison group.

Cohort 2 Students' Results

Assessments at school entry showed that the Intervention, Intervention^{+}, and Comparison groups were functionally similar in terms of foundational reading-related skills.

At the end of Year 1, the Intervention group outperformed the Comparison group on almost all variables, including phonological awareness, alphabetic coding, language processing, word reading and spelling.

On the last assessment occasion for Cohort 2 at the middle of Year 2 (mid-2017), the Intervention students outperformed the Comparison students on all three reading and spelling assessments. Of particular importance was the finding that the low decile Intervention group was superior to the low decile Comparison group on almost all assessments. Mean scores for low decile Comparison students on some variables approached or equalled those of higher decile students.

There were no statistically significant differences between the groups for reading self-efficacy. Scores were generally high (positive). However, it is perplexing that the highest, most positive, scores were obtained for the low decile Comparison group, which consistently had the lowest scores on almost all variables.

Our expectation that Cohort 1 Intervention teachers might consolidate PLD learning and achieve better results with another group of students (Intervention^{+}) was not fulfilled. The Intervention^{+} students did not achieve significantly better results than the Comparison students on most variables.

An important finding was that the Cohort 2 Intervention group significantly outperformed the Cohort 1 intervention group on middle of Year 2 reading and spelling assessments.

Cohort 2 Teachers' Results

Teacher surveys and observations focussed on Intervention teachers. Results were available for 20 teachers who participated in the 2016 PLD workshops, and who had pre- and post-intervention survey data.

Teacher knowledge of language foundations associated with literacy teaching and learning was moderate prior to the intervention, with an average of 61% of knowledge questions answered correctly. There was a significant increase to an average of 75% correct answers following the PLD workshops.

Teachers' self-evaluations of their knowledge was more positive for the comprehension aspects of teaching reading prior to the intervention. The post-intervention self-evaluation revealed greater confidence in using phonics in balance with comprehension instruction.

Data on teaching practice showed a change from instruction that incidentally taught students the alphabetic code, to practice that was more explicit in teaching the code. Teachers generally showed greater attention to the use of word-level decoding skills and strategies.

The teachers' coach observed that all teachers used a commercial phonics programme. This observation is consistent with separate findings that 85%-90% of New Zealand schools use a phonics programme. Most teachers also used commercially produced alphabet cards. Assessment of students' foundational language-related reading skills appeared to be *ad hoc.* Teachers tended to continue reliance on context-based cues when students encountered an unfamiliar word in text, although video observations revealed an increase in word-level cues.

Overall, Cohort 2 Intervention teachers made important changes in their knowledge, instructional practices, and self-confidence. These changes appear to be reflected in the improved literacy learning outcomes for their students.

Answers to the Research Questions

Answers to the research questions for Cohort 1 and Cohort 2 are summarised in blue font adjacent to each question:

1. Did students in the intervention groups show improved literacy learning outcomes compared with students in the comparison groups?
    - Cohort 1: No.
    - Cohort 2: Yes. The Intervention group significantly outperformed the Comparison and Intervention^{+} groups at the end of Year 1, and in the middle of Year 2.
2. Did students in the intervention groups show improved motivation in reading compared to those in the comparison group?
    - Cohort 1: No
    - Cohort 2: No. Reading self-efficacy, as a proxy for reading motivation, was similar across all groups in both cohorts. Scores were generally positive. Self-efficacy scores were somewhat independent of actual reading-related achievement, suggesting that the scale we used was inadequate.
3. Did the intervention reduce the literacy achievement gap?
    - Cohort 1: No
    - Cohort 2: Yes. Low decile Intervention students achieved results that were markedly better than those of students in the Comparison group. In many cases, low decile students performed at levels that approached or were close to students in higher decile schools.
4. Did the intervention result in increased teacher confidence in teaching word-level skills?
    - Cohort 1: No

- Cohort 2: Yes. Results indicated improved levels of teacher knowledge as well as more positive self-perceptions of ability to teach phonic-related knowledge and strategies.

Conclusions

PLD programmes designed to modify instructional strategies to enhance literacy learning outcomes require a well-resourced approach that explicitly and systematically supports teachers with the implementation of new teaching methods. Meaningful changes were observed in Cohort 2 teachers' instructional behaviours following significant modifications to the PLD workshops and teacher supports.

The significantly improved literacy learning outcomes for Cohort 2 Intervention students were consistent with the theoretical basis of the research project. Initial word learning, which forms the basis of reading comprehension, is dependent on students quickly mastering the ability to decode words in text. Many students benefit from systematic and explicit instruction in word-level decoding and the language foundations associated with the alphabetic principle.

For over three decades, focus on explicit word-level decoding strategies has not been favoured in New Zealand schools or supported by instructional materials and students' readers. Yet, international and New Zealand research clearly indicates that changes to literacy instruction which include a greater emphasis on systematic and explicit instruction aimed at students developing efficient word decoding skills can have a major, positive impact on literacy learning outcomes. Findings from this research project provide further support for the importance of such a change in literacy teaching. Students in low decile schools derived important benefits from this shift in teaching emphasis.

Despite challenges associated with changing deeply embedded instructional practices, the project demonstrated that modifications can occur in literacy instructional practices through a research-led PLD programme based on a strong theoretical framework.

Teachers and students throughout New Zealand stand to gain if the findings from this research project are used to form the basis of a nation-wide strategy for changing significant aspects of literacy instruction in New Entrant/Year 1 classrooms. We encourage the Ministry of Education to use the findings from this research to develop such a strategy.

TABLE OF CONTENTS

INTRODUCTION

What is this Project About?

The purpose of the Early Literacy Project was to examine the effect on students' literacy learning outcomes of a teacher professional learning and development (PLD) programme designed to improve the literacy learning outcomes of students, especially those in lower decile schools. The teacher PLD programme was developed to provide teachers with the knowledge and skills to adopt explicit and systematic word-level decoding strategies and skills in their literacy instruction with New Entrant/Year 1 students. The approach was based on research indicating that many students, and especially those in low decile schools and from diverse backgrounds, are likely to achieve better literacy learning outcomes if literacy instruction is more explicit and focussed on the development of word-level decoding skills and associated language skills that underpin reading acquisition and reading comprehension.

The Project participants comprised two groups of teachers of New Entrant/Year 1 students, and two cohorts of New Entrant/Year 1 students. The first group of teachers took part in PLD workshops during 2015. The second group participated in PLD workshops during 2016.

We collected reading and reading-related data on students who were in these teachers' classrooms on a number of occasions during the course of the Project. *Cohort 1* students commenced school in February 2015. Data were collected over a 2½ year period from school entry through to the middle of 2017 when the students were in Year 3. *Cohort 2* students commenced school in February 2016, with data collected from school entry through to the middle of 2017 when the students were in Year 2. Furthermore, we collected data from students in schools that chose not to participate in the PLD workshops but which were willing to serve as points of comparison.

We also collected a range of data from teachers who participated in the project. Assessments were undertaken on teacher knowledge of the language constructs associated with early literacy acquisition, as well as teachers' self-perceptions of their ability to effectively teach students to learn to read. In addition, we examined teachers' word identification prompts and conducted video observations of a sample of each teacher's literacy lessons.

In this report we present information on the rationale for the research; the nature of teachers' professional learning and development workshops; the nature of the student and teacher assessments; results from the assessments; conclusions about the impact of the research on student literacy learning outcomes and on the effectiveness of the teacher PLD programme; and, finally, implications and recommendations for policy and practice.

Why Was this Research Undertaken?

Concerns have been expressed about the literacy performance levels of New Zealand children and adults for close to 20 years. The concerns arose from the results of international surveys of students and adults, as well as from data collected by the Ministry of Education (MoE). It became evident during the 1990s that, compared to many OECD countries, New Zealand had the largest spread of scores between good and poor readers (Elley, 1992), and that low-performing readers were more likely to be Māori and/or from low-income backgrounds (Wagemaker, 1993). Research conducted 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 advantaged and disadvantaged 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, 1997).

Concerns about the literacy skills of New Zealand children continued through the first decade of the 21st Century. National Standards were introduced in 2010 by the government as one means to regularly identify children's progress in literacy and numeracy. In the year following the introduction of National Standards, the MoE's Briefing to the incoming Minister of Education (Ministry of Education, 2011) following the 2011 general election, identified some overall improvement in education but noted that disparities in learning outcomes appear early and often persist throughout learning. The Briefing concluded that "The greatest challenge facing the schooling sector is producing equitable outcomes for students" (p. 23).

The most recent international survey of Year 5 students' literacy performance (PIRLS: Progress in International Reading and Literacy Survey; Mullis et al., 2017) revealed that New Zealand students' literacy performance has declined in comparison to other countries, despite various initiatives over 17 years to improve literacy learning outcomes.

The research in this project is based on a substantial body of theory and results from many previous studies in New Zealand and overseas. The Project was designed to test the view that students benefit from an approach to literacy instruction that places greater emphasis on the development of literacy-related language skills and word-level identification strategies.

What Does Contemporary Research Say About Literacy Learning?

Scientific research on how students learn to read indicates that achievement in reading comprehension performance depends on the ability to recognise the words in 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, or mappings, provides the basis for constructing the detailed orthographic representations required for the automatisation of word recognition (or what Ehri, 2005, calls *sight word knowledge*). When this occurs, cognitive resources can be allocated to sentence comprehension and text integration processes (Pressley, 2006), that is, to the *meaning* of text.

To discover mappings between spelling patterns and sound patterns, students must also be able to segment spoken words into subcomponents. Students 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). Understanding the alphabetic principle, or "cracking" the alphabet code, is necessary (but not sufficient by itself) for being able to read for meaning.

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 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).

There is now 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 covering 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 students and crucial for some.

HOW CAN CURRENT APPROACHES TO EARLY LITERACY INSTRUCTION BE ENHANCED?

Building Teacher Knowledge to Increase Word-level Instruction

A key focus of this research project was on PLD workshops for teachers. The goal of these workshops was to provide additional teaching strategies designed to improve the literacy learning outcomes of students.

Effectively teaching reading skills to beginning readers requires that teachers have a high level of understanding of the basic structure of the English language, including an understanding of the sound–symbol correspondences of written English and how these influence reading development. In her seminal publication, *Teaching Reading IS Rocket Science,* Moats (1999) showed that relevant teacher knowledge includes understanding language and linguistics, and being well versed in the psychology of reading development and in effective instructional practices that are based on research-informed models of cognitive development.

More specifically, the National Reading Panel in the United States (2000) undertook a comprehensive review of research on reading acquisition. The Panel identified five "big ideas" central to success in early reading: *phonemic awareness*, which is the ability to hear and manipulate the sounds in words; *alphabetic principle*, which is the ability to associate sounds with letters and use these sounds to form words; *accuracy and fluency*, which is the effortless, automatic ability to read words in connected text; *vocabulary*, which is the ability to understand and use words to obtain and express meaning; and, *comprehension*, which is the essence of reading involving the intentional interaction between reader and text to derive meaning. All five elements are required for success in reading.

Before teachers are able to effectively teach students to read or to develop the foundation skills for learning to read, it is important that they are not only knowledgeable about the code of written and spoken English, but also have knowledge of research-based literacy assessment and instructional procedures (Gersten, Compton, Connor, Dimino, Santoro, Linan-Thompson & Tilly, 2008; Spear-Swerling & Zibulsky, 2014).

Piasta, Connor, Fishman, and Morrison (2009) have highlighted the role of teacher knowledge. They found that the development of word-level decoding skills was crucial for successful reading acquisition. However, time spent on explicit decoding instruction was only effective for student word-learning growth when teacher knowledge of phonology, orthography, morphology, literacy acquisition, and instruction was high.

To adequately meet the needs of all students, knowledge of contemporary, research-based effective literacy practices must also be part of a teacher's toolbox for literacy instruction. Yet, because of the different theoretical understandings of what reading is and how it should be taught, many teachers have not had the opportunity to refresh their teaching skills in line with contemporary scientific research on best practice in literacy instruction. However, even when teachers have sufficient contemporary *knowledge* of appropriate instructional practices, they often struggle to implement or plan for them in their lessons (McNeill & Kirk, 2014; Spear-Swerling & Zibulsky, 2014). McNeill and Kirk (2014), for example, found that for the teaching of spelling, teachers were generally familiar with a variety of evidence-based practices, but tended not to use them because they felt that they lacked the explicit knowledge of how to adopt them in practice. Additionally, Fielding-Barnsley (2010) found that pre-service teachers in both early childhood and primary education programmes knew the importance of teaching phonic knowledge to beginning readers, but they lacked a clear understanding of how to use such knowledge in an explicit and systematic instructional approach.

Professional Development in Literacy Teaching

In this research project, the focus of our teacher PLD programme was on developing in teachers a high level of the teacher knowledge that is required for effective teaching. Teacher knowledge of English orthography and morphology makes it easier for teachers to understand the patterns for word decoding and word spelling, and in turn, this understanding helps them to assist students in learning the essential skills for reading and spelling (McNeill & Kirk, 2014). Students who do not acquire an understanding of the patterns, either through explicit teaching or implicit learning, start to lag behind in their literacy development. They become reliant on identifying unfamiliar words in text by guessing or using non-text cues (e.g., illustrations), strategies which characterise poor

readers (Nicholson, 1991, 1993; Pressley, 2006). Thus, the focus was shared between linguistic content knowledge and strategies for the explicit teaching of that knowledge to beginning readers.

A Framework for the Teaching of Word-Level Knowledge

We adopted the Cognitive Foundations of Learning to Read framework (Tunmer & Hoover, 2014; see Figure 1) in this project. This framework combines the cognitive elements underpinning the development of the language comprehension and word recognition components of the Simple View of Reading (Gough & Tunmer, 1986). It is based on research showing that learning to read follows a developmental progression from pre-reader to skilled reader involving qualitatively different but overlapping phases. Skill in comprehending written text depends on the ability to recognize the words of text accurately and quickly; 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 the ability to discover mappings between spelling patterns and sound patterns in turn depends on letter knowledge, phonemic awareness, and knowledge of the alphabetic principle (e.g., Ehri, 2014). Language comprehension is also necessary, particularly in terms of vocabulary, morphological knowledge, as well as syntactic knowledge (Tong, Deacon, & Cain, 2014).

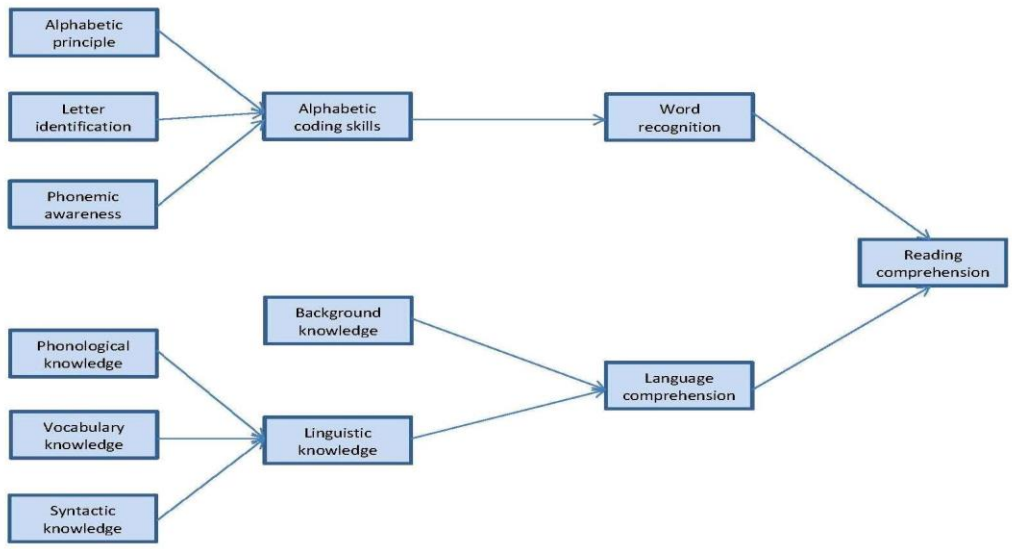


Figure 1: Cognitive Foundations of Learning to Read (Tunmer & Hoover, 2014)

The literacy learning needs of beginning readers necessarily vary because they differ in the amount of reading-related knowledge, skills, and experiences they bring to the

classroom on school entry, 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.

Progress in learning to read, in both word recognition and language comprehension, is dependent on the child-by-instruction interactions that occur during instruction (Arrow & Tunmer, 2012). Child-by-instruction interactions can be described using Byrne's (2005) *division of labour for acts of learning* framework that takes into account differences students bring to the process of learning to read. Within this framework, the division of labour assumes that any act of learning is a product of both the environment and the learner. Byrne (2005) argued that different acts of learning could be located along different points of the continuum representing the division of labour between the learner and the environment. At one end of the continuum, acts of learning require less structured and often fragmentary environmental input for learning to occur (such as learning spoken language), whereas the other end of the continuum represents learning that requires rich and highly structured input from the environment (such as learning calculus).

For some beginning readers, the processes of acquiring literacy skills are highly learner-dependent because some students grasp the idea of what is needed to discover orthographic patterns after relatively small amounts of explicit teaching of phonologically-based skills and strategies. Other students, however, are more environment-dependent, and benefit much more from a fairly structured and teacher-supported introduction to reading. At school entry, learner-dependent students typically come from more advantaged backgrounds and bring with them higher levels of essential reading-related knowledge. On the other hand, environment-dependent students tend to come from less advantaged backgrounds and have more limited amounts of essential reading-related knowledge. Therefore, differentiated teaching, where teachers use evidence-based assessment procedures and instructional strategies, can cater to the different literacy learning needs of beginning readers from the outset of schooling.

The structure of the Cognitive Foundations framework provides the basis for diagnostic reading assessment. For example, if beginning readers are not progressing satisfactorily in learning to derive meaning from print (i.e., reading comprehension), it is because they are having problems understanding the language being read (i.e., language

comprehension), problems recognizing the words of text quickly and accurately (i.e., word recognition), or both. Weakness in word recognition skill stems from insufficient explicit instruction in alphabetic coding skills or inadequate opportunities to practice and receive feedback on applying alphabetic coding skills while actively engaged in reading. If alphabetic coding skills are still weak despite exposure to explicit instruction and practice, it is because of inadequate knowledge of the alphabetic principle, letter knowledge, or phonemic awareness.

The structure of the framework does not imply that the development of the more advanced cognitive elements cannot occur until all of the more fundamental elements are fully developed. Although some level of mastery of the more fundamental elements of the framework is needed to develop mastery of the more advanced ones, the elements tend to develop congruently and reinforce each other in a reciprocally facilitating manner. The elements of the framework should therefore not be taught in isolation from each other but instead should be taught in an integrated manner; beginning readers should be given plenty of opportunities to practice and receive feedback on applying their newly acquired skills while engaged in performing the more advanced cognitive functions specified in the model. A literacy programme should offer students opportunities to be immersed in language development, including immersion and engagement with children's literature.

Explicit Instruction

The second component of the PLD programme was the provision of strategies for teaching the knowledge and skills. This specifically revolved around the use of explicit, structured, and systematic instruction. Explicit instruction is a form of instruction that requires the breaking down of the learning required into observable elements.

In early literacy instruction the elements include each letter name and sound, blending of phonemes together to decode unfamiliar words in their entirety, segmenting sounds in spoken words to spell unknown words, as well as teaching the meanings of target words. As each of the fundamental elements are taught they are built on by increasingly complex elements, such as consonant and vowel digraphs, syllables and morphemes, and extending into syntax and punctuation. Ehri's (2014) phase model of word recognition provides a framework for understanding how the increasingly complex elements are used to develop word recognition fluency and automaticity.

Explicit instruction also necessitates a systematic approach to teaching that ensures that the intrinsic cognitive load is minimised during a teaching sequence. This means that the teaching sequence itself is structured so that the learners don't need to think of anything except the very specific element that they are learning at any one time. To do so, a lesson sequence is carefully structured in the same way each time. This includes a revision of the last element taught, the explicit and direct instruction of what the new element is to be learned, followed by the teacher modelling the learning sequence and how to use the new element in reading or spelling.

This explicit instruction is followed by students practising the new element as the teacher works with them. In the earliest learning sequences, this typically involves the students being taught at least one consonant and vowel sound by the teacher telling students what the letter is and what the sound is, and students following along with saying the name and sound, generally with magnetic letters or through forming the letters on a whiteboard. This sequence is followed by practising blending the new elements and any previously learned elements to read and create new words. The key to this sequence is the final part in which students practice the new elements within text; the text however, must only contain the new and previously learned elements. As there is no part of the lesson that is not previously or currently taught the learner need not put cognitive effort into identifying what they are learning.

In Short

The Cognitive Foundations framework, combined with Ehri's (2014) phase model, can be used to provide guidance for teachers in recognising the developmental progression in learning to read. Teachers must be able to recognise the developmental progression to make instructional decisions that take into account students' existing reading-related knowledge. Knowledge of the specific phonic elements required for effective word learning is also necessary. It is our contention that the teacher knowledge required for this is low for New Entrant/Year 1 teachers in New Zealand. This, along with a history of multiple-cue, constructivist approaches, means that many teachers are unable to provide explicit phonic instruction in their reading programme even with the help of commercial phonics programmes.

The lack of explicit and effective phonic instruction in the new entrant classroom is less problematic for children who come to school with already high levels of reading-related knowledge. They will continue to learn to read by building on this existing body of knowledge. Students, predominantly in low-decile schools, who have low levels of reading-related knowledge, benefit from explicit teaching of the phonic elements required to transition from being a non-reader to a beginning reader. By increasing the phonic knowledge of teachers and providing explicit guidance in the developmental progression, we anticipate that teachers will be better equipped to meet the needs of all learners, not just learner-dependent students. In particular, we anticipate that teachers who engage in teaching phonic knowledge in explicit and systematic ways to students in low-decile schools will result in students achieving reading and spelling outcomes that are closer to those of students in middle and high decile schools.

WHAT WERE THE MAIN RESEARCH QUESTIONS IN THIS PROJECT?

Following from our extensive reviews of relevant research literature before and during the course of this project, a number of research questions have been raised and tested in earlier research reports.

The focus of our research questions is on the effect of the teacher PLD programme on the literacy learning outcomes of New Entrant/Year 1 students in their class during participation in the programme, and in subsequent years. We were interested also in the effect on students' motivational factors. In addition, to make assumptions about the ability for students to show improved literacy learning outcomes relative to students whose teachers did not participate in the PLD programme, we posed research questions regarding teacher knowledge and evidence that those teachers who participated in the PLD programme made changes to their literacy teaching practice.

For this final research report, we address the following research questions, as requested by Ministry of Education officials (letter of 6 September 2017):

1. Will students in the intervention groups show improved literacy learning outcomes compared to students in the comparison groups?
2. Will students in the intervention groups show improved motivation in reading compared to those in the comparison groups?
3. Will the intervention reduce the literacy achievement gap?
4. Will the intervention result in increased teacher confidence in teaching word-level skills?

These research questions are considered separately for Cohort 1 and Cohort 2.

COHORT 1 (FEBRUARY 2015 – JULY 2017)

This longitudinal research project involved two cohorts of teachers and students. In this section, we report information about and findings for Cohort 1, which covered the 2½ year period from February 2015 to July 2017.

How Did We Select Schools for Participation in the Project?

This project started at the beginning of 2015. Based on Ministry of Education data for the 2014 school year, we estimated that it would be possible to identify approximately 1,600 New Entrant students in schools throughout the lower North Island. It was our goal to have around 800 students in an Intervention group and the same number in a Comparison group. With this in mind, a random sample of schools was selected for participation in the project from regions of the lower North Island that include Wellington, Hutt Valley, Wairarapa, Kāpiti, Horowhenua, Manawatū, Whanganui, Ruapehu, Tararua and Taranaki.

A stratified frame was used in an attempt to maximise participation of lower decile schools in the project. The initial selection process was drawn from state and integrated primary schools listed on a Ministry of Education database. Included in the draw were 80 schools that were expected to enrol eight or more New Entrant students at the start of 2015.

We excluded schools from the Rangitīkei and Ruapehu districts because of their small size and the small number of New Entrant children expected to be enrolled at the start of the 2015 school year. In addition, to avoid confusion and "cross-contamination", we excluded schools in the Porirua area and parts of Wellington and Hutt Valley that were participating in the *Shine Literacy Success for All* project.

Schools were randomly selected and randomly allocated to either the Intervention or Comparison conditions. This procedure was performed by means of a random number generator in the SPSS statistical package. The names of the schools were not known until the random selection process was completed.

Following the sampling process, schools were identified and principals of the selected schools were contacted and invited to participate in the project and to attend meetings in Wellington, Palmerston North, and Whanganui to discuss the goals and activities of the project. Explanations included which of the two groups, Intervention or

Comparison, to which the school had been randomly assigned. Ministry of Education School Liaison staff assisted with the recruitment process.

Principals of the 80 schools drawn randomly from a Ministry of Education database were contacted towards the end 2014 seeking agreement to participate in the project. By the start of the 2015 school year, less than half the schools approached (39) confirmed their willingness to take part in the research.

Our goal was to have a randomised control research design. However, because considerably fewer schools than we approached agreed to participate in the study, the result is a quasi-random volunteer sampling design.

What Were the Characteristics of the Teachers and Students?

Of the 39 schools that agreed to participate in the project, 24 had been randomly assigned to the intervention group, and 15 to the comparison group. A total of 62 teachers of New Entrant students were identified: 38 in the intervention schools and 24 in the comparison schools. These numbers fluctuated as teachers came and went for various personal or professional reasons. At the first of the scheduled intervention group teacher professional development workshops, 45 teachers attended from the 24 intervention schools.

Time 1 baseline assessment data were collected during February and early March 2015 from 359 New Entrant students. Of these, 201 (56%) were in intervention schools, and 158 (44%) were in comparison schools.

The mean age of the sample at the time of first assessment was 60.56 months, which is around 5 years, 6 months; the median age was 60 months, and the mode was 60 months. Clearly, the large majority of students were around 5 years of age. There was no statistically significant difference between the mean ages for the Intervention and Comparison groups: 60.69 and 60.38 months respectively.

In terms of gender, 52% of the students were boys and 48% were girls. However, there was a marginal imbalance for students in the intervention group: 54% (107) were boys compared to 46% (93) girls. The gender breakdown in the comparison group was even: 51% (80) boys and 49% (78) girls.

Regarding ethnic background, information was available for 312 (87%) of the sample. New Zealand European/Pakeha students comprised 63.8% (199) of the sample; Māori were 24.7% (77); Pasifika were 5.1% (16); Asians were 3.5% (11); and Others were 2.9% (9). Some schools were unable to provide ethnic background information for all students.

Decile rankings of schools showed some differences between the intervention and comparison schools. We grouped decile rankings as follows: low = deciles 1-3; medium = deciles 4-7; high = deciles 8-10. The spread across these three decile bands was slightly more even for the Comparison schools than the Intervention schools. These data are presented in Table 1.

Table 1: Distribution of project participants by group and school decile band.

	Group	
	Intervention	Comparison
Decile band	Percent	Percent
1-3	26.4	36.7
4-7	54.3	31.3
8-10	19.2	31.9

What Assessments and Procedures Did We Adopt?

We obtained assessment data from students on seven occasions. In this report we focus on data collected at the start and end of Year 1 (2015), end of Year 2 (2016), and middle of Year 3 (2017). Our goal here was to determine whether students in the Intervention and Comparison groups were similar in terms of literacy-related skills and knowledge at the start of 2015, and whether the students in the Intervention group showed any improvement in literacy learning outcomes over the course of the project when compared to students in the Comparison group. We also wanted to compare Intervention teachers' knowledge of literacy factors and instruction with the Comparison teachers' knowledge early in Year 1 and towards the end of the year.

Assessments with Students

We employed research assistants to administer the assessments to students. All assistants followed the same procedures for each individual assessment but were free to provide the assessments in the order that worked best for individual students. When an assessment appeared overwhelming for students they returned to their classroom and completed that assessment at a different time, either that day or on another day. Students were assessed in a quiet break-out space near their classroom, or in another quiet space in the school. Some students were assessed in a quiet corner of the classroom when other students were there. Research assistants were not informed as to which were Intervention or Comparison schools.

The following assessments were undertaken during February 2015: letter name and letter sound knowledge; receptive vocabulary knowledge; word recognition; invented spelling; and phonological processing. Details of these assessments are presented in Appendix 1.

Teacher Survey

In addition to the student assessments, we carried out a survey of teachers' literacy knowledge, self-evaluations of their literacy knowledge, literacy teaching self-efficacy, and teacher word identification prompt scenarios. These results were presented in a separate report to the Ministry of Education (Chapman, Arrow, Tunmer & Greaney, 2015). A summary of information included in that report is presented in appropriate sections of this report. Specific details about each of these teacher assessments are presented in Appendix 2.

What Did We Do in the Teacher PLD Workshops?

We ran five teacher PLD workshops during the year in Wellington, Palmerston North and Hawera. The first workshop, in March, was conducted over two days. The other four workshops were for one day each in May, July, September and November.

The PLD programme was designed to provide research-based strategies for teachers to supplement instruction in their existing literacy programmes. Teachers in the

Intervention group were asked to attend all five of the workshops during the course of 2015 on how to teach word-level skills to beginning readers, within language-orientated approaches. Between the workshops, an online interactive forum site was developed to enable Intervention teachers to share ideas and strategies.

The PLD programme comprised five modules, each with some core components that occurred in each module. These components covered the content knowledge for teaching (e.g., different vowel sounds and their spelling patterns, the explicit strategies of sounding out and blending for decoding); analysing assessment data in terms of students' knowledge of the content and their strategy knowledge; and, ways of teaching content knowledge to students, including planning instruction using lesson planning templates that were provided in the workshops. Assessment data collected during the previous assessment occasion were used by the facilitators for discussion in the workshops.

Teachers were requested to access the online community and website where forums, video clips, word document templates of lesson plans and other useful resources were provided. Teachers were also expected to review their video observations, and to provide a reflection using the confidential conversation tool on the website. The PLD facilitators monitored the online community to provide support and guidance as required. Unfortunately, most teachers did not make use of the website as expected, and all but two or three teachers were resistant to providing reflections following viewing of their videos.

The PLD modules were developed to correspond with the developmental nature of reading, as illustrated in the Cognitive Foundations Framework (Figure 1). The content of the first four modules was specifically linked to each corresponding element of the framework (vocabulary and phonological awareness; alphabetic principle; alphabetic coding; linguistic comprehension). The final module drew together the content of the previous modules to show how differentiated instruction can be implemented in the classroom. This module also drew on the participants' experiences in applying the teaching approaches covered in the previous modules. Details for each module are presented in Appendix 3.

WHAT WERE THE RESULTS FOR TIME 1 ASSESSMENTS?

Student Assessment Data

Group Comparisons

We adopted a Group (Intervention vs. Comparison) by Decile Band (Low, 1-3; Mid, 4-7; High, 8-10) analysis of variance design. An initial multivariate analysis of variance (MANOVA) was performed with 10 Time 1 variables. These variables were as follows:

- Letter name, upper and lower case
- Letter sound, upper and lower case
- Clay word test
- Clay word phonemes
- Invented spelling
- Invented spelling sounds
- CTOPP Elision
- CTOPP Blending
- CTOPP Matching
- Receptive vocabulary, BPVS (British Picture Vocabulary Scale, standard scores)

Overall, the two groups were similar in their performance on these Time 1 assessments; there were no statistically significant differences between the groups. However, an examination of means revealed that for seven of the 10 variables, the Comparison group obtained higher scores; for two variables, the Intervention group obtained slightly higher scores, and for the BPVS the standard scores were almost identical. Means and standard deviations are presented in Appendix 4a.

There were statistically significant differences in terms of Decile Band. Overall, the High decile group obtained higher scores, followed by the Mid decile group, with the Low decile group obtaining lower scores. Statistically significant Decile Band differences were observed for Letter Sound, CTTOP Elision, CTTOP Matching and BPVS. Results for these analyses are presented in Appendix 4b.

To summarise, the Intervention and Comparison groups were reasonably similar at Time 1, although the Comparison group generally had higher scores than the Intervention group. These higher scores were not statistically significant. In terms of decile bands, students in low decile schools tended to obtain significantly lower scores on these reading-related assessments than students in mid and high decile schools.

Teacher Survey Data

Fifty-five responses were received from the teacher survey conducted online through the *Survey Monkey* platform. Despite requests to complete all sections of the survey, not all of the 55 responses included complete data for all sections. Fewer responses were received for the teacher word identification prompt questions, which were located in the last section of the survey, than for the teacher efficacy questions at the start of the survey.

Teacher Knowledge

Data from the teacher knowledge of basic language constructs of literacy survey were analysed in terms of the types of teacher knowledge and as a function of Intervention and Comparison groups. For the self-evaluation of literacy teaching knowledge, 54 valid responses were analysed by means of a *t*-test. As expected, there was no significant difference in mean scores between teachers in the Intervention and Comparison groups. Mean scores for each of the eight self-perceptions of literacy teaching-related skills were mainly in the "moderate" to "very good" categories (over 90% of responses). The only area in which there was less perceived skills related to teaching English language learners: 23% thought they had "minimal" knowledge for working with such students.

For the teacher skills and knowledge of language constructs, total scores were calculated for phonemic, phonic, phonological and morphological variables. There were no statistically significant differences between teachers in the Intervention and Comparison groups.

Results in terms of the percentage of questions answered correctly in each section are as follows. For phonemic knowledge/skills, the Intervention teachers

answered on average 62% of these items correctly; for the Comparison teachers, an average of 69% of the phonemic items was correctly answered. For the phonic knowledge/skills items, 52% of Intervention teachers' responses were correct, and 57% of Comparison teachers answered these items correctly. Phonological knowledge/skills items were generally answered correctly: 89% for Intervention teachers and 90% for Comparison teachers. Morphological knowledge/skills, however, were less well understood: 52% correct for Intervention teachers and 54% correct for Comparison teachers.

Teacher Prompts for Word Identification Errors

Teacher prompts for each of six reading error scenarios were scored following Greaney's (2001) approach.

Overall, 40% of the prompts were word-level. These included such cues as "Let's sound that word again"; "can you see two words?"; "hear and say all the sounds you see"; "what comes after *p*... and then *d*... that rhymes with *dad*?"; "look at the blend at the start and try again"; "Let's see if looking at the chunks in the word can help".

On average, 45% of the prompts were based on context. Examples included: "Try that again and think what would make sense"; "Look at the picture then try again"; "Go back to the start of the sentence and think what will fit"; "Think about the story, what would make sense"; "Does the word you read match the picture?"

Neutral prompts accounted for an average of 15% of the cues teachers reported using. These cues were generally lacking in useful information for helping the reader: "Try that again"; "That was lovely, but I wonder if you can find your mistake?"; "You made a mistake. Can you find it? Fix it?"; "Have a go"; "Good job. Good reading"; "Check it"; "Get your mouth ready".

In general, initial and total word level responses for the word identification scenarios tend to be fewer than 50% of the prompts teachers reported favouring. However, Intervention teachers showed greater use of word-level prompts than Comparison teachers. This emerging difference may have been due to information provided during the first PLD seminar.

Teacher Efficacy

Total scores for the Literacy Teaching Efficacy Scale were analysed to examine whether there were differences between the Intervention and Comparison teachers in terms of confidence to bring about a range of literacy outcomes in the classroom. The results revealed comparable and generally positive means for the two groups: Intervention *M* = 228.57 (*SD* = 39.16), Comparison *M* = 221.21 (*SD* = 38.86). The slight difference in means was not statistically significant.

Intercorrelations

Pearson product-moment correlations were computed for the teacher knowledge, self-evaluation, and teacher efficacy variables. There was a clear disjunction between teachers' self-evaluation of literacy-related knowledge and the measures of linguistic knowledge. Correlations for the self-evaluation variable ranged from a high of .34 with phonological knowledge to a low of .15 with phonemic knowledge. Perhaps not surprisingly, self-evaluations correlated reasonably highly with teacher efficacy: *r* = .53. However, all of the teacher efficacy correlations with the teacher knowledge variables were very low, ranging from -.01 (phonemic knowledge) to .16 (morphological knowledge). These results suggest that teachers generally hold fairly high levels of self-evaluation and teaching efficacy, but that these are not reflected in their levels of knowledge for key language constructs associated with literacy teaching and learning.

Teacher Comments About the Survey

Respondents were provided with the opportunity to comment on the nature of the survey; 22 teachers provided comments. Six responses were negative, and included comments such as "too long", "we teach New Entrants…this is expecting us to be linguistic experts", and "too much pressure for busy NE teachers". Other teachers were positive and grateful for having the opportunity to participate in the survey: "It was really hard but the challenge was great because it really made me think about my knowledge and how I can apply it"; "Thanks for making me think!"; "Thank you, this survey really got me reflecting upon and analysing the strategies I am using during guided reading sessions and in class generally". Some comments were more mixed: "It was hard! Highlights things I don't know

and maybe should know and using"; "I wish I hadn't sat down to do this late at night! Interesting to reflect on though"; "Some very tricky questions! Some I had no idea about".

WHAT DO THESE COHORT 1 BASELINE RESULTS MEAN?

Student Results

Overall, findings from the school selection and randomisation process, together with data from the Time 1 assessments, indicate that students in the Intervention and Comparison groups were generally similar in terms of age and gender. More students were in the intervention group (56%) compared to the comparison group (44%). In terms of decile bands, fewer Intervention than Comparison students were from low decile schools, and very few Comparison students were from middle decile schools. Similar percentages of Intervention and Comparison students were enrolled in high decile schools.

The standardised scores for receptive vocabulary were very similar for both groups of students. It is also worth noting however, that although not statistically significant, the Intervention group had slightly lower overall school entrance literacy abilities than the Comparison group.

An examination of the distribution characteristics of scores for almost all measures (the exception was receptive vocabulary) revealed strongly skewed scores. On many variables, large numbers of students in both groups scored at "floor" levels. This finding is normal and expected for young students who have just entered school.

Finally, as expected, differences as a function of school decile band were marked in these Time 1 data. The differences were especially notable for receptive vocabulary knowledge, letter knowledge, and measures of phonological processing, with low decile students performing at lower levels than high and middle decile students.

Teacher Survey Results

The overall purpose of the Teacher Survey was to provide information about literacy knowledge and efficacy for teachers participating in the project. Not all teachers responded to all items in the survey. The response attrition rate may have been due to the length of the survey. Some commented that the survey was too long. Some items were not answered, especially in the teacher knowledge of language constructs section. Comments from several teachers suggested that it was not appropriate for them to know about aspects of literacy-related language constructs that formed the teacher knowledge survey.

This viewpoint is disappointing but not surprising considering the lack of emphasis on language constructs in teacher education over the past four to five decades.

Although both groups of teachers had comparable levels of knowledge across the four language constructs domains, there were very high levels of understanding of phonological skills/knowledge, medium levels of phonemic skills/knowledge, but lower levels of phonic and morphological skills/knowledge. Interestingly, although the scores revealed strengths in the area of phonological awareness, only 58% of teachers were able to provide an accurate definition of phonological awareness. This finding is similar to that reported by Washburn et al. (2011), and suggests that phonological knowledge is incomplete.

The questions associated with the alphabetic principle/phonics knowledge were more difficult for teachers. The accuracy rate for this section of the teacher knowledge survey was only 54%. Effective literacy instruction has been consistently shown to include systematic teaching of phonics (e.g., Adams, 1990; National Reading Panel, 2000). Accordingly, explicit knowledge of phonics principles is required for teaching decoding and spelling (Washburn al., 2011). It is concerning that only around half of the teachers in this survey were able to correctly identify when to use key, reliable phonics principles.

Aspects of morphology were the most challenging for teachers, with an overall accuracy rate of 52%. These findings are somewhat consistent with those reported by Moats (1994), who found that graduate level teachers had considerable difficulty with various aspects of morphology.

In general, teachers had a mixed understanding of the literacy-related language structures required for effective teaching. As Mather et al. (2001) commented, teachers with insufficient grasp of such crucial language structures are unlikely to effectively teach reading skills explicitly to those students who show early signs of developing reading difficulties.

The data on teacher prompts from the six reading error scenarios showed that overall, fewer than 50% of the first prompts were word-level cues. In general, context and neutral cues were used together more frequently by teachers. This preference reflects the advice presented in publications on literacy teaching for beginning readers (e.g., *Reading in Junior Classes; The Learner as a Reader*; *Effective Literacy Practice in Years 1 to 4*).

Results of the teacher efficacy scale showed generally high levels of self-efficacy in regard to a range of literacy teaching situations. Relatively high self-ratings are common for

such scales. Although teachers reported positive levels of confidence in their literacy teaching abilities, this confidence did not relate to their actual knowledge of key language constructs associated with literacy learning.

A similar pattern of results was found with teachers' self-evaluation of their literacy teaching skills; the overall responses indicated that teachers felt they had moderate to very good levels of literacy teaching skills. Intervention and Comparison teachers were comparable in their self- evaluations.

The finding that teachers held high self-efficacy and high self-evaluations for the teaching of different components of reading, when their actual knowledge was low, is in line with existing research findings (Cunningham, Perry, Stanovich, & Stanovich, 2004; Spear-Swerling, Brucker, & Alfano, 2005). Classroom teachers are not generally aware of the knowledge that they don't know and thus feel confident in their abilities and knowledge.

WHAT DID WE FIND AT THE END OF YEAR 1 FOR COHORT 1?

Student Sample

A total of 304 students remained in the sample at the end of Year 1. Compared to the 359 students at the start of the project, the difference represents an attrition rate of 15.3%. Of the students who remained in the study at the end of Year 1, 178 were in the Intervention group (attrition rate = 11.3%) and 126 were in the Comparison group (attrition rate = 20.3%; this higher attrition rate was due to one school withdrawing from the project part way in 2015).

Not all students had complete data, which resulted in a further reduction in the sample. Despite numerous attempts to obtain complete data for all students in the project, some schools were unable or unwilling to meet our requests.

Of particular importance was the effect of students' being moved into different classrooms in regard to the intervention sample. At the start of the project, we explained to schools participating in the Intervention that it was important for students to remain with the same teacher throughout the year. Our explanation included the crucial fact that teachers who participated in all of the teacher professional learning and development workshops throughout the year would have the greatest potential to improve the literacy learning outcomes of the New Entrant/Year 1 students.

This request was not met by a number of schools. As a result, some Intervention group students had different teachers during Year 1, some of whom had participated in the PLD workshops and some who had not. We discuss the effect of these changes in the results section.

Student Assessments

The same procedures for collecting the assessment data were used as for Time 1 at the beginning of the year. Some assessments, such as letter knowledge, were not used at the end of the school year because typically students reach ceiling on those measures by that time. Additional assessments were introduced to match the developmental progression expected of students following completion of a year's literacy learning.

The following assessments were administered during November 2015:

- invented spelling; phonological processing;
- word identification;
- pseudoword reading;
- reading book level.

Details about each of these assessments are presented in Appendix 5.

WHAT RESULTS DID WE OBTAIN AT THE END OF YEAR 1?

Student Data

Effects of Teacher Changes During Year 1

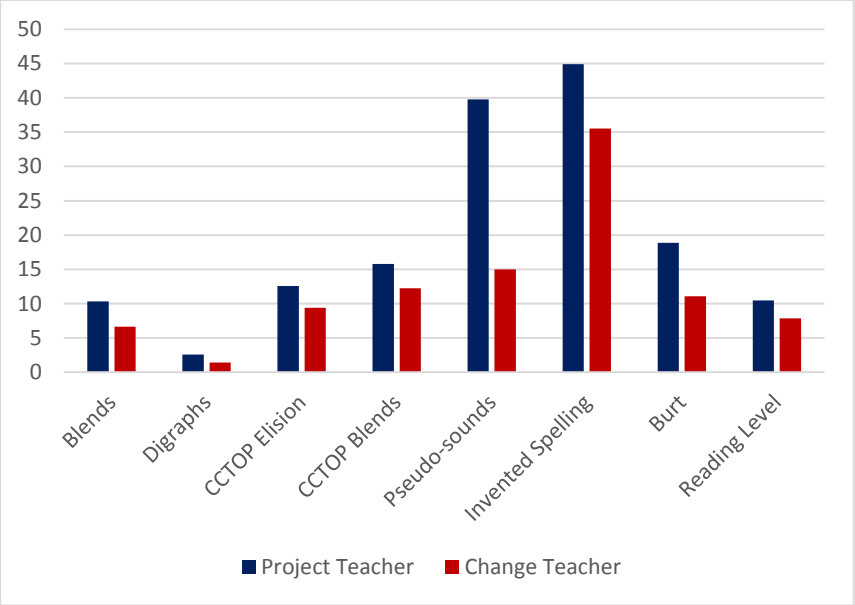
As we indicated earlier, a number of schools were unable to meet our request that students in the Intervention have the same New Entrant/Year 1 teacher throughout the year. This has had an important effect on the results and also on the sample size. Of the 24 Intervention schools, six schools made teacher changes for Intervention students that involved moving them to other participating teachers. Modern learning environments also made it difficult to maintain 'line of sight' from teacher to student. In this cohort we had two schools that had team-teaching or modern learning environment teachers. Teachers in these schools were also participating in the project PLD workshops and working with Intervention students. Seven schools moved Intervention students into classes with teachers who were not part of the project, and two schools had project teachers who were absent for significant periods of time during the year.

Overall, forty-five Intervention students had new teachers who were not part of the project, or project teachers who were absent for lengthy periods of time during the year. We analysed end of Year 1 assessment data to determine whether major teacher changes (i.e., Intervention students having a non-project or absent teacher during the year) differentiated those students from other Intervention students who either had the same teacher throughout the year or a different teacher who was involved in the project PLD workshops. The results of these analyses are important.

All of the eight end-of-year assessments showed highly statistically significant mean score differences between Intervention students who had the same or another project teacher during the year compared to students who had a different or "absent" project teacher during the year. In each case, the students with project teachers outperformed those who had significant teacher changes. The results reveal that skills emphasised in the PLD sessions (especially phonological processing as measured by Pseudoword sounds and Invented spelling sounds), along with word knowledge (Burt word test), resulted in markedly higher scores for Intervention students whose teachers remained consistent

throughout the year. These differences are illustrated in Figure 2, and summary data are presented in a table in Appendix 6.

Figure 2. **End of Year 1 Intervention Group Results by Teacher.**



Note: The y axis scale has been removed because of the different metrics for each measure. The graph is designed simply to illustrate comparisons between student outcomes as a function of having the same or different teacher during the first year of the study.

Because of these marked effects, we eliminated those students with major teacher changes from further analyses of Cohort 1 group comparison data.

This decision resulted in a reduction of the Intervention sample from 200 to 155. Further reductions occurred as a result of missing or incomplete data. As was the case in regard to baseline assessment data, not all students or schools completed or supplied assessment data. Especially disappointing was the incomplete Reading Book Level data.

The factors associated with these reductions in sample size were unsatisfactory and compromised the initial desire to obtain as large a sample as possible with Intervention teachers working with their New Entrant students for the entire first year.

Results of Analyses of Student Data

Given the significant effect that project versus non-project teachers had for Intervention students, we analysed end of Year data only for students who had a project teacher throughout the year, and compared them with the Comparison students. We

adopted the same analysis design used for the school entry data; that is a Group (Intervention with project teacher vs. Comparison) by School Decile Band design.

The two key research questions to be addressed by the end of Year 1 data were whether the Intervention group started to outperform the Comparison group, and in particular whether low decile students in the Intervention group were starting to outperform low decile students in the Comparison group.

Process Variables

First, we conducted a multivariate analysis of variance (MANOVA) on the following "process" variables: blends, digraphs, pseudoword phonemes, CTTOP elision, CTTOP blending, CTTOP matching, and invented spelling sounds. The results revealed that the Comparison group obtained higher scores on the Digraphs variable than the Intervention group. There was also a significant effect for Decile band; students in low decile schools performed at lower levels than those in mid and high decile schools for all of these process variables.

Outcome Variables

We also conducted analyses on "outcome" measures: Burt word reading, reading book level (where these were available), spelling, and pseudoword reading. The Comparison group obtained significantly higher scores than the Intervention group for spelling and pseudoword reading. As with the process variables, there was also a significant effect for Decile band; low decile students obtained lower scores on all variables than mid and high decile students. We also found that the High Decile Comparison group obtained higher scores for Reading Book Level than the High Decile Intervention group.

Summary data for these end of Year 1 results are presented in Appendix 7a and 7b.

Correlations

Finally, in regard to end of Year 1 data, we computed product moment correlations between school entry variables and end of year reading outcome variables (viz., Burt word test, Reading Book Level) to identify key predictors. Included in the entry variables were receptive vocabulary knowledge, Letter Name and Letter Sound knowledge, Invented Spelling Sounds, the phonological processing variables of elision, blending, and matching,

and Clay Word Phonemes. The correlations were calculated for the complete sample of students who had scores for the relevant variables.

The strongest predictor of the Burt word test was Letter Sound knowledge (combined upper and lower case scores), *r* = .68, followed by Letter Name knowledge (combined upper and lower case scores), *r* = .66.

For Reading Book Level, the highest correlation was Letter Name Knowledge, *r* = .63, followed by Letter Sound Knowledge, *r* = .61.

These data are consistent with other studies indicating that alphabet knowledge is highly predictive of later reading ability.

Interestingly, receptive vocabulary knowledge was only moderately predictive of Burt word scores (*r* = .42) and Reading Book Level (*r* = .38). Correlations are presented in Appendix 8.

Teacher Data

Teacher Knowledge

To determine whether any changes occurred in teacher knowledge of literacy-related language constructs and self-evaluation of literacy between earlier in the year and end of year, a series of analyses were performed.[1] Disappointingly, data were available for only 24 teachers who had both sets of assessment scores (Intervention *n* = 14; Comparison *n* = 10). None of the domains of teacher knowledge resulted in the Intervention teachers outperforming the Comparison teachers at the end of Year 1. It appeared that the Intervention teachers were unable to expand their knowledge of basic language constructs involved in literacy teaching and learning as a result of the workshops. The result was disappointing.

We also performed product moment correlations between teacher self-evaluations and knowledge of language constructs. None of the correlations between self-evaluations and each of the knowledge variables was statistically significant; the highest correlation was .14 (phonological knowledge). This finding suggests that teachers' relatively positive self-evaluations of their literacy-related teaching knowledge do not relate to their actual knowledge of basic language constructs.

[1] We used analyses of variance with repeated measures for these analyses.

Teacher Video Data

Teacher video data showed teacher practice during small group instruction. There were four major findings from examination of teacher practice observations in relation to the content of the workshops.

1. A reliance on the *Ready to Read* book series as the curriculum *and* process for teaching reading.
2. A reliance on embedded implicit teaching and an associated lack of systematic, explicit teaching focused on a child or group's developmental progression and needs.
3. Some teaching of word knowledge (alphabet, sounds, blends) but no teaching students to apply this knowledge in the reading process.
4. A dominance of teaching strategies that demand high input from the learner, rather than teacher input towards a gradual release of responsibility.

Implications of these Teacher Video Data

These observations are consistent with the lack of improvement shown by Intervention teachers as a result of the PLD workshops. They provide insights as to why Intervention students did not start to show improvements in literacy learning outcomes when compared to the Comparison students.

WHAT DID WE FIND AT THE END OF YEAR 2?

Cohort 1 Sample Description at the End of Year 2

At the start of the project in 2015, Cohort 1 comprised 359 students. At the end of Year 2, 264 Cohort 1 students remained in the project: 156 (59%) were in the Intervention group and 108 (41%) were in the Comparison group. The Intervention group had 58% males and 42% females, whereas the Comparison group was more evenly matched with 51% males and 49% females.

Regarding ethnic background, data from schools were still unavailable for 20 students. Based on the information that was supplied, the following percentages were in each group:

- Intervention: Pākehā = 63%; Māori = 24%; Pasifika = 9%; Other = 5%
- Comparison: Pākehā = 62%; Māori = 27%; Pasifika = 4%; Other = 7%

In terms of decile bands of participating schools, the following distribution was observed:

- Intervention: low decile = 23%; middle decile = 58%; high decile = 19%
- Comparison: low decile = 31%; middle decile = 44%; high decile = 25%

The Assessments We Used

Assessments used with Cohort 1 students at the end of Year 2 were clustered as follows:

Foundations of Literacy: CTOPP Elision; pseudoword reading phonemes.
Word Recognition: Burt word test; contextual reading accuracy; pseudoword reading; spelling accuracy.
Language Comprehension: Listening comprehension; mispronunciation task.
Reading comprehension and fluency: Neale reading comprehension test, Neale fluency test.
Details of these end of Year 2 assessments are presented in Appendix 9.

Students' Results

Our primary analysis design for treating the end of Year 2 data was by means a series of Group by Decile Band (2 x 3) multivariate analyses of variance (MANOVAs). Each MANOVA was performed in terms of the following clusters of variables:

- *Foundations of literacy*
- *Word knowledge*
- *Language comprehension*
- *Reading comprehension and fluency*

The analyses revealed statistically significant group differences for foundations of literacy, word knowledge and language comprehension; the Comparison group outperformed the Intervention group. For Reading comprehension and fluency, the differences between the two groups were not statistically significant.

For Reading book level, we performed a univariate ANOVA. The result revealed that the Comparison group (M = 18.46) was marginally better than the Intervention group (M = 17.64), but this difference was not statistically significant.

Regarding Decile Band, analyses revealed that the low decile group performed at lower levels on most variables than the middle and high decile groups.

Means and standard deviations for the end of Year 2 assessments are presented in Appendix 10a, and results for the analyses of variance are presented in Appendix 10b.

These results were very unexpected. Based on earlier research (e.g., Tunmer, Chapman, & Prochnow, 2003), we had anticipated that any ground work in terms of the literacy foundations that may have been taught on the basis of information gained from the PLD might have taken longer to influence the literacy outcomes of Intervention students. This was not the case.

WHAT DID WE FIND AT THE MIDDLE OF YEAR 3?

Cohort Sample Description at the Middle of Year 3

We noted that at the end of year 2, 264 Cohort 1 students remained in the project: 156 (59%) were in the Intervention group and 108 (41%) were in the Comparison group. At the middle of year 3, which was the final data collection point, 231 Cohort 1 students were assessed by research assistants: 128 (55%) were in the Intervention group and 111 (45%) were in the Comparison group. The apparent increase in the Comparison group is due to some participants in this group being absent at the end of Year 2 during data collection. In terms of gender, the Intervention group had 58% males and 42% females; the Comparison group had 52% males and 48% females.

Data regarding ethnic background were available for all but eight students across the Intervention and Comparison groups. The following numbers were in each group:

- Intervention: Pākehā = 76 (66%); Māori = 23 (20%); Pasifika = 7 (6%); Other = 9 (8%)
- Comparison: Pākehā = 70 (65%); Māori = 27 (25%); Pasifika = 3 (3%); Other = 8 (7%)

In terms of decile bands of participating schools, the following distribution was observed:

- Intervention: low decile = 27 (22%); middle decile = 69 (58%); high decile = 24 (20%)
- Comparison: low decile = 38 (34%); middle decile = 44 (40%); high decile = 29 (26%)

Consistent with previous experience, we were unable to obtain all reading book level data for all students in the project; of the 231 students who were in the project at the middle of year 3, 18 reading book levels were not supplied by schools. Other data were missing as the result of students being absent from school at the times assessments were being undertaken.

The Assessments We Used

Assessments used with cohort 1 students at the middle of Year 3 were clustered as follows:

Word Recognition: Burt word test, contextual reading accuracy, spelling accuracy.

Language Comprehension: Listening comprehension, mispronunciation task, British Picture Vocabulary Scale, Morpheme test (described in Appendix 11).

Reading Comprehension and Fluency: Neale reading comprehension test, Neale fluency test, Reading Book Level (where available).

Word Attack Strategy (described in Appendix 11)

Reading Self-Efficacy (described in Appendix 11)

Students' Results

Given the results reported for the end of year 2 data, we did not anticipate any marked change in the differences between the Intervention and Comparison groups.

Following from the previous analysis design, each Group by Decile Band MANOVA was conducted on the cluster of variables identified in the previous section.

Unlike the results for the end of Year 2, none of the MANOVAs resulted in statistically significant Group differences.

The separate ANOVA performed on Reading book level resulted in a non-significant Group effect; the Intervention group obtained a slightly higher mean than the Comparison group (20.01 vs. 19.50).

Similarly, the ANOVA performed on Reading Self-Efficacy also resulted in non-significant difference between the Intervention and Comparison groups.

Regarding Word Attack Strategy, we examined the two response options (word-level vs. text level) in terms of percentages within each group. While 80% of students in the Intervention group responded that they used a word-analysis approach for identifying unfamiliar words in text, 75% of students in the Comparison group made this response. Chi-square tests revealed that these differences were not statistically significant.

In regard to Decile band, across most variables the students in the low decile band tended to perform at lower levels than those in the middle and high decile bands.

Results for Reading Book Level revealed that Low Decile Intervention students obtaining higher book levels than the low decile Comparison students.

Results for Reading Self-Esteem were not statistically significant in terms of Decile Band.

Means and standard deviations for the Cohort 1 Middle of Year 3 assessments are presented in Appendix 12a. Data from the statistical analyses are presented in Appendix 12b.

Summary

Overall, the students' results at the middle of Year 3 were disappointing, but not unexpected given the trends that were occurring on previous assessments.

WHAT DID WE LEARN FROM THE COHORT 1 RESULTS?

Anecdotal Observations and Teacher Feedback from the PLD Sessions

The teacher workshops provided the opportunity to obtain feedback from teachers about the implementation of the teaching strategies and materials.

Many teachers were supportive of the project and the content of the PLD sessions. Teachers indicated that they were developing an awareness of how to plan a scope and sequence for literacy instruction based primarily on Ehri's (2005; 2014) phases of developing word recognition. Many teachers reported using the phases to plan the instruction for students, rather than providing incidental instruction derived from the emphasis on text and meaning.

Numerous teachers indicated that focusing more on the skills associated with word recognition meant that they needed to slow down the initial progress through book levels to ensure students had a more comprehensive grasp of the foundation literacy knowledge. With this change in emphasis, teachers reported that their students were beginning to analyse print before resorting to using picture cues, or guessing unknown words in text.

In addition, teachers often commented that they felt that the pressure of getting through a book was reduced, particularly for students who had difficulty accessing most of the text in the book. 'Slowing down to speed up' and 'working through the word' became two common themes revealed in the workshops.

However, teachers also indicated that they had difficulties in implementing the scope and sequence and the more focused word-level instruction. In terms of our target students, many teachers noted several times that it was too late for them to make effective changes for the Intervention students based on the PLD sessions that ran throughout the year.

This difficulty with applying significant changes to literacy instruction arose because teachers' new knowledge was developing at the same time. Implementation of instructional changes was clearly a challenge for many teachers. In some cases, changes began during the middle of the year for the Intervention students; in other cases, teachers were reluctant to make major instructional changes during the course of the year.

Other issues that arose were more systemic in nature. Many teachers voiced concerns that it was difficult to reconcile the current approach to literacy instruction in New Zealand schools with the approach presented in the PLD workshops. Combining the current approach, which emphasised deriving meaning from text, with the project materials, which placed a greater emphasis on teaching word-level decoding strategies, was clearly challenging.

There was a common view among the Intervention teachers that the levelled book series made it difficult for them to focus more on phonic knowledge and the related strategy of blending. Furthermore, many teachers indicated that they struggled with the developmental progression of phases alongside the current emphasis on text content, rather than developing strategy knowledge. Requirements to 'fill the gaps' to meet schools' expectations about students' reading book levels were also presented as major difficulties in adopting the project materials and strategies.

In some cases, teachers reported that they were not permitted in their school to use magenta texts for more than four weeks, or they were not allowed to give students a yellow level text until they had 'passed' the red 3 running record.

Despite these significant challenges it was notable that many teachers in the final workshop commented that it was in students' writing where they saw the greatest changes. This may be understandable if the only place that teachers felt they had more flexibility for implementing phonic knowledge is in writing, where it is completely necessary.

Based on the various teacher data we concluded that overall, many teachers were unable to make significant changes to their knowledge or to their practices.

In the PLD workshops, most teachers readily engaged with new learning about how students' word knowledge develops progressively, the importance of helping students develop orthographic maps, and the act of teaching explicitly.

The teacher videos, however, revealed that teachers had difficulty in applying this learning to their small group instructional time, which in most cases is dominated by the use of a selected text as the way to teach reading.

Use of Phonics

Teacher interviews indicated that many relied on a published phonics programme for teaching word knowledge. It appears from the observations that often this knowledge is

isolated to a phonics time and not specifically applied during small group reading or writing instructional sessions. It is during small group instruction that teachers can direct such teaching to the specific needs of individual students, and demonstrate the strategies that will help them in reading a particular text.

The similar levels of phonological awareness and initial alphabetic coding knowledge, such as blends and digraph knowledge, across the Intervention and Comparison groups can be explained in part by the use of phonics programmes in many schools (see Chapman, Greaney, Arrow, & Tunmer, 2017).

Feedback from teachers during the PLD sessions is consistent with these observations and provides compelling evidence for the failure to observe trends in the right direction for the Implementation students.

Instructional Changes and the Influence of Reading Texts

The primary goal in our study was for Intervention students to progress further in developing the language skills shown by research to be necessary for enhancing word identification strategies during the first year of reading instruction.

However, we did not expect to observe the difficulties that teachers had in translating the knowledge and trial practice from the PLD sessions into different instructional approaches in the classroom.

One reason appears to be the incompatibility between the stronger word-level decoding approach to initial reading instruction that is emphasized in the PLD sessions on the one hand, and the deeply embedded teacher practice of using the existing book series as the vehicle for teaching reading.

These texts are structured to teach reading via a process approach, such as a reliance on meaning, structure and possibly the first letter of an unknown word. The teacher observation data showed most teachers used the texts in this manner. Teacher support materials with the *Ready to Read* series promote this process approach.

The reading series is generally based on a teaching paradigm not structured to suit a developmental approach to the learning and teaching of word knowledge (Ehri, 2014). Rather, the series is framed by the assumption that students learn to read largely by exposure to words in text.

This is clearly not the case for students who struggle with learning to read. The texts expose students to words from across a large number of word patterns. This approach counteracts efforts to systematically teach word-level decoding skills.

Our data suggest that a reliance on these texts for teaching students to read appears to have left both teachers and students confused when teachers attempted to implement practice based on the PLD workshops. This confusion likely contributed to the poorer performance of the Intervention students in relation to the Comparison students, because teachers were unable to overcome embedded practices.

For example, in one lesson the teacher was attempting to help students decode a red level text that included the words *want*, *what* and *does*. While sentence structure could carry some of the students' success with such a text, it is not useful as a text or a process for helping students to develop increasing levels of automaticity in word recognition. The range in orthographic patterns of those three words did not allow for an appropriate decoding strategy.

We hold that it is highly significant that on the one hand, teachers see the need for teaching students about word level concepts (as seen by the number of phonics programmes added to current practice in most schools), and the need for explicit and systematic teaching. But on the other hand, most teachers have been trained in and are familiar with a book series approach *that requires a different type of teaching*. In essence, the two approaches are incompatible.

The video data, together with the teacher knowledge survey results, provide a plausible explanation for why the Intervention students did not start to outperform the Comparison students during this first year of the project.

Reflections on the Delivery of Teacher PLD

As a result of the unexpected and disappointing student achievement results for the Intervention group at the end of Year 1, together with our reflections on the teacher observation and survey data, we concluded that changes should be made to the delivery and emphasis in PLD workshops with a new cohort of teachers and students.

We saw the need to include a range of more explicit instructional strategies with appropriate "scope and sequence" elements to facilitate earlier embedding of word decoding instruction in teachers' regular classroom literacy activities. A much more explicit

approach in demonstrating the very specific ways in which teachers can significantly increase their emphasis on the development of word level skills was clearly necessary.

In keeping with the tendency in New Zealand for teacher PLD to take place within a co-constructivist framework, we were less explicit and directive than we should have been. We made assumptions that teachers would be able to apply their new knowledge as a result of this co-constructivist approach.

However, an examination of the relevant teacher implementation literature indicates that teacher PLD involving different pedagogical approaches than those currently practised by teachers should provide both explicit content knowledge and a practical 'how-to' guide for putting this new knowledge into practice. This more explicit approach was considered by us to be more effective than our initial approach in the PLD sessions of leaving teachers to explore and trial implementation on their own (Desimone, et al., 2002; Garat et al., 2001; Pianta et al., 2008; Stahl et al., 2013).

COHORT 2: (FEBRUARY 2016 – JULY 2017)

WHY DID WE INTRODUCE A SECOND COHORT INTO THE PROJECT?

Introduction

The results presented in the first part of this report were clearly disappointing. We have discussed factors that likely contributed to the failure of the PLD to have a positive effect on students' literacy learning outcomes when compared to students in the Comparison group.

Among the various reasons is the manner in which the teacher PLD sessions were structured and delivered. It became apparent to us that changes needed to be made to the nature and delivery of the PLD if the overall goals of the project were to be realised.

During the latter part of 2015 and early 2016 we redesigned the PLD workshops and reconceptualised the nature of the delivery. We approached a number of schools, seeking participation as part of a second cohort for the project.

Changes to PLD Programme and Workshops

Changes to the PLD programme in 2016 were made to provide more systematic guidance on how to implement the "new" instructional strategies into effective practice. Also, the delivery of PLD workshops was changed to provide for much more explicit demonstration of the very specific teaching strategies designed to foster the development of word level skills.

Why didn't we adopt this approach in the first place? With Cohort 1, we made the false assumption that when teachers engage in PLD workshops, they prefer to work in a co-constructivist framework with minimal direct, explicit guidance. This assumption was false, as we learned from teachers' feedback. We also assumed that the teachers would be able to adapt their practice more easily than they did. We under-estimated the impact of the core levelled text (*Ready to Read*), the embedded practice of guided reading and its incidental approach to teaching word knowledge, and the use of running records on levelled texts as the way of measuring a student's progress.

For Cohort 2, we explicitly emphasised the research-based view that phonics instruction provides a 'kick-start' to phonological decoding for students who come to

reading with few of the necessary cognitive entry skills; that is, those students who have little understanding of the interaction between the graphemes of printed words, and phonemes of spoken words (Tunmer & Greaney, 2010).

Such provision of both explicit content knowledge and a 'how-to' guide for putting into practice rather has resulted in more effective outcomes for students (Desimone et al., 2002; Garet, Porter, Desimone, Birman, & Yoon, 2001; Pianta et al., 2008; Stahl, Keane, & Simic, 2013). We initially chose not to base our PLD workshops along those lines because of the prevailing approach in New Zealand to teacher PLD. However, our decision to make the change to more explicit guidance is supported by research such as that just cited.

Recent New Zealand research on literacy has also found that explicit and direct teacher guidance during PLD workshops, and in developing scope and sequence routines, does contribute to improved outcomes for students, particularly those in low decile schools (e.g., *Shine Literacy Project*, Chapman, Allcock & Cody, 2016; *Quick 60 project*, Chapman, 2016).

Accordingly, we developed an explicit 'how-to' guide: The *Early Literacy Project Curriculum* (ELPC). This guide provided a scope-and-sequence framework within which teachers could identify the learning needs of their students within one of Ehri's (2014) four developmental phases.

Ehri's phases were specifically adapted for this PLD programme to facilitate a focus on the foundation skills that most students need for the first six months of formal education, and that some students need for a longer period of time.

In the ELP Curriculum, a sequence for content instruction (e.g., letter-sounds, blends, digraphs, syllables, suffixes) was recommended for each phase. The necessary instructional strategies were placed alongside the content so that teachers could readily identify how students might more effectively learn how to read *and* spell words containing those elements.

The ELP Curriculum document also provided teachers with a sequence of high-frequency words for their students to learn "by sight". Guidance was provided on how to teach the elements for the whole class, as well as how to engage in more personalised differentiated instruction.

The ELP Curriculum also provided lists of existing resources, programmes, and text types that are often used in New Zealand classrooms, and how they align with students' levels of reading development. Alongside this scope-and-sequence was information on what to look for in assessment data to best align students' progress with the phases of reading development.

A lesson plan template that illustrated how teachers make instructional decisions based on specific needs supplemented the Curriculum. We also provided recommendations on how to choose reading materials that were aligned to the use of explicit instruction.

Table 2 (following two pages) presents a summary of the content and strategy knowledge required at each phase. The ELP Curriculum itself provides a more specific weekly sequence for teaching the content of each phase.

One of the challenges in making the shift in instructional approach relates to the nature of resources available to schools. Many of the beginning books currently available to schools contain a myriad of spelling patterns and words that students may not have previously come across, *in one book*. These books require teachers to focus on meaning-based cues, because the words themselves are too difficult for many students to use phonic-based strategies during the early phases.

Simply said, existing resources and book series (e.g., *Ready to Read*, *PM+* readers) are not aligned with the teaching strategies that are required for effective learning during the earlier phases of reading development.

Some schools that had sufficient literacy funds purchased more appropriate resources, but others were unable to. Additional materials were purchased by the Project to assist schools with inadequate funds to acquire resources that were consistent with research-led content in the PLD sessions. These materials were provided to support the teaching of letter-sound mappings, blends and digraph content knowledge, and blending strategies. Also included were decodable readers that enabled the reinforcement and practice of letter-sound, blend or digraph patterns that had been taught. Decodable texts are specifically structured so that only previously taught letter-sounds, blends, digraphs or high-frequency words were encountered by the beginning reader.

Table 2: Summary of phases, teaching activities and aligned resources

Group: Phase 1	**Group: Phase 2**	**Group: Phase 3**	**Group: Phase 4**
• Alphabet names & sounds: see in print and words • Syllable and rime awareness • 1:1 matching • Concepts about print • Attempts at writing and spelling	• Letter sounds (vowels) • Consonant blends • High frequency sight words • Basic punctuation • Spelling letters for sounds • Decoding, using blends and paying attention to all letters • Blending and segmenting phonemes	• Trigraphs & vowel structures • Syllable spelling patterns • Morphemes • Advanced Word attack using chunks/or unitizing for decoding instead of sounding out • Spelling using chunks including morphemes • Cross-checking across meaning and syntax for decoding attempts • Checking for meaning	• Syllable structures • Morphological knowledge including role of meaning • Comprehension strategies • Syntactic structures • Analogy for decoding unfamiliar words • Morphemes for identifying meaning of unfamiliar words • Purpose of reading beyond learning to read
Explicit teaching activities			
• Explicit syllable and rime awareness instruction • Matching spoken word to printed unit reinforced with predictable texts • Initial sound sorting • Teaching names and sounds	• Explicit instruction in blends, sounds, sight words and strategy use • Reinforce with decodable text • Sounding and blending • Segmenting and blending • Say it and move it • Irregular and regular high frequency words	• Explicit instruction in patterns, strategy use, sight words and cross-checking attempts • Reinforce with a combination of decodable and levelled text • Irregular and regular high words • Analogy use • Teaching letter patterns • Question clusters • Direct comprehension instruction	• Explicit instruction reinforced with levelled text • Question clusters • Direct comprehension instruction • Story mapping • Text structure • Summarisation

Independent extension and home activities			
• Alphabet ring cards with letters that are already learned and being learned attached. • Mum and Dad encouraged to read child's library book to them and find letters in it.	• Decodable texts or teacher created sentences to re-read • Alphabet & high-freq words on rings to practice for fluency	• Texts with taught units within – levelled texts • Spelling words with taught units (not tested, just practiced) • Whisper reading	• Texts with taught units within • Spelling words with taught units (not tested, just practiced) • Silent reading • Asking questions about story and discussing
Centre/rotation activities			
• Handwriting using letters • Using phonic apps to practice sounds of letters • Dictating stories to teacher, peer, or into book apps • Reading known picture books and shared reading books – finding known letters	• Creating words with letters • Sentence construction with h-f words and punctuation cards • Handwriting using letters and words • Partner reading decodable texts • Phonics apps	• Games or apps to reinforce larger units • Handwriting and spelling using learned units • Genre writing using learned units and sight words • Partner reading	• Games or apps to reinforce larger units • Handwriting and spelling using learned units • Genre writing using learned units and/or summarise story • Partner or silent reading
Resource types to use			
• Letters (magnetic, plastic, foam, flash cards) • Letter-sound flip-charts (e.g., Smart Kids) • Smart Kids phonics 1 • Yolanda Soryl stage 1 resources • Predictable texts • Alphabet books • Sounds like Fun (Allcock)	• Decodable texts (Yolanda Soryl EW readers 1 & 2; Word detective; Letterland; About Words; Jolly Readers; Little Learners) • Alphabet cards & resources • High-frequency word cards • Punctuation cards • Smart Phonics 2 kit • Yolanda Soryl stages 3-5 • Sounds like Fun	• Levelled texts (yellow +; Word detective) • Syllable and morpheme apps or games • Smart Phonics 3 kit • Yolanda Soryl stages 6-7 (if available)	• Levelled texts (Green +; all trade books including Ready to Read) • Comprehension activities and resources • Syllable and morpheme apps or games

These additional resources were sent to schools based on the level of students' need at the earlier phases of development where such resources are most necessary and beneficial. In this way, those schools which had students with less need during the earlier phases were sent fewer resources; those schools with students who had greater need were sent more. The resources were combinations of:

- Magnetic letter sets (consonants, vowels, blends, consonant and vowel digraphs, word families), and small decodable readers that students can keep. These were sourced from *Smart Kids*, based in Auckland;
- A set of decodable texts (controlled phonic-knowledge introduction texts) sourced from *Little Learners Love Literacy*, based in Australia;
- A small set of phonic-knowledge emphasis texts from the *Word-Level Readers* sourced from Gilt-Edge publishers, based in Wellington.

Workshop Programme

Although the approach taken within the workshops was changed, the content focus remained the same.

The PLD modules were developed to correspond with the developmental nature of reading, as illustrated in the Cognitive Foundations framework (see Figure 1). The content of the first four modules was specifically linked to the elements of the framework:

- Vocabulary and phonological awareness
- Alphabetic principle
- Alphabetic coding
- Linguistic comprehension

The final module drew together the content of the previous modules to show how differentiated instruction can be implemented in the classroom. This module also drew on the participants' experiences in applying the teaching approaches covered in the previous modules. The timeline and focus content of each workshop is presented below.

1. **March 2016.** Introduction, understanding the research and theory for the project, explicit teaching, initial teacher knowledge building in phonological awareness and phonic concepts (digraphs, vowel patterns, morphemes);
2. **May 2016.** Initial phonic teaching (consonants and simple vowels, blends and digraphs), integrating content and strategy use for reading, explicit vocabulary teaching;
3. **August 2016.** Later phonic teaching (vowel digraphs, syllable types, early morphemes) and integrating the content and strategy use into text reading;
4. **September 2016.** Language comprehension for reading comprehension including syntax, punctuation, early strategy use;
5. **November 2016.** Lessons learned and using the learning for differentiated instruction.

Workshop Modules

Module 1: Introduction and the importance of language

In this module teachers were introduced to the cognitive development of reading framework, and the associated assessment framework. This first module included an introduction to effective instruction, including the roles of direct, explicit instruction and implicit learning at different phases of reading development.

The emphasis in this section was on the need for explicit, systematic and structured instruction for students who are most at risk of having difficulties in learning to read. In contrast, for students not at risk, more implicit instruction is just as useful as they seldom require explicit instruction in the foundational knowledge.

A key element of effective instruction is to distinguish between the linguistic content students need to learn for reading and spelling unfamiliar words, and the strategies they need to learn to make use of that content, as well as the strategies they need to learn to ensure accuracy and fluency.

The linguistic content, and how students learn to use that content, can be captured for word recognition using Ehri's (2014) phases of word recognition and learning. To support instructional decision-making within the phases, the module also included a session on using a range of assessments for screening, diagnostic purposes and for monitoring progress. An

assessment booklet with samples of assessment tools and an explanation of assessment was also provided.

Module 2: Understanding letter knowledge and phonological awareness: learning how to read words

In this module teachers were introduced to the role of vocabulary in decoding and language comprehension, as well as an introduction to the nature of phonological awareness.

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 (Senechal, Ouellette, & Rodney, 2006; Tunmer & Chapman, 2012a; Tunmer & Chapman, 2012b).

Students 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 feature as major influences on reading comprehension because the inability to recognize the words in text limits the ability to understand text. However, this does not suggest that instruction in foundation skills should be delayed until students have acquired fast, accurate word recognition skills (Tunmer & Chapman, 2012b). The module focused on explicit teaching of selected words and their meanings within an authentic story used during interactive oral reading.

The specific developmental processes of letter-sound knowledge and their relationship with phonological awareness, was covered in the second half of this module. The way that they interact to contribute to alphabetic coding skills was emphasised.

This content captured the main learning requirements of the first two phases in the ELP Curriculum. A large body of scientific research indicates that comprehending text in an alphabetic orthography depends on the ability to recognize the words in 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).

In this module teachers were provided with content knowledge for distinguishing between vowels and consonants, how the sounds are similar and how they differ, as well as how students make use of sounding out and blending for learning to read words independently.

Module 3: Developing word knowledge for fluency

In this module teachers were introduced to different word reading strategies that students need to learn, and how they are used in conjunction with each other.

The teachers were given specific instructions in the different long vowel sounds and digraphs, distinguishing between blend sounds and digraphs, and identifying morphemes and different syllable types in words. The content in this module covered the learning required in phases 2 through to 4 in the ELP Curriculum.

Once students reach this point of development, explicit instruction is seldom needed for further word recognition and decoding. In the ELP Curriculum, this aligns with the end of phase 3 and into phase 4. It is expected that students add to their knowledge of print and engage in self-teaching from their initial foundation store of words.

By this point students should be fluent readers who are building vocabulary and reading comprehension by way of deeper understandings of text. They can now competently attempt to identify unfamiliar words of varying complexity without needing instruction in all unfamiliar print words prior to reading a text.

Module 4: Reading comprehension as the goal

This module covered language comprehension beyond vocabulary as well as explicit reading comprehension strategy instruction and its place in year 1 classrooms. Teachers were provided with instruction on ways to teach text structure as a form of background knowledge (part of the Cognitive Foundations Model, see Figure 1). In addition the module covered sentence construction (de Jong & van der Leij, 2002; Ouellette & Beers, 2010; Tunmer & Chapman, 2012b) including parts of speech, sentence complexity and punctuation. Initial comprehension instruction for beginning readers is less directed than word reading and vocabulary building but is still necessary as low levels of knowledge in these areas contribute to later reading comprehension difficulties (Tong et al., 2014).

Module 5: Differentiated instruction as the goal

Ideas for reconceptualising how to use whole class and small group instruction for the differentiated classroom from the start of the school year were the focus for this module.

Such changes had been introduced from Module 1 through the use of the templates guiding teachers to rethink their small group and whole class instruction, including not only the *how*, but also the *what was taught*.

A large part of this module involved revisiting the previous modules and recapping how the phases work together and what they mean for practice. During this module, teachers also shared their experiences, provided feedback and completed the teacher knowledge survey.

Teacher Coaching

In 2015 we anticipated using the online website as the place where we would provide coaching, answer questions, and communicate with teachers who were in the workshop programme. However, engagement with the online forum was minimal. Although there were over 1000 views throughout 2015, there were less than 10 posts that reflected levels of interaction with workshop materials or classroom teaching experiences.

We attempted the online coaching approach again at the beginning of 2016, but we observed even less success. To overcome this lack of interaction, we appointed a "coach" to engage in face-to-face support. The coach was a retired but experienced literacy teacher and teacher educator, who had previously held the role of Resource Teacher of Reading, and as such was familiar with students' specific literacy needs. This person visited most of the Cohort 2 2016 workshop teachers on four occasions.

COHORT 2: WHO PARTICIPATED?

Selection of Cohort 2 Schools

We obtained agreement from 34 teachers in 13 schools to form a new Cohort 2 Intervention group starting in 2016. All schools had been in the Comparison group in 2015.

We also obtained agreement from the 20 schools which were part of the 2015 Cohort 1 Intervention sample to have their New Entrant/Year 1 teachers (n = 40) continue in the project with 2016 New Entrant students. We refer to this group as "Intervention^{+}". This step was taken to examine whether participation in the 2015 workshops was associated with improved literacy learning outcomes for "new" students, based on the assumption that these teachers had had an opportunity to reflect on and consolidate their knowledge from the PLD they received in 2015.

We formed a new Cohort 2 Comparison group from five schools. One school had been a Comparison school in 2015 and chose to continue as a Comparison school. The remaining schools had been in the project during 2015, but the teachers in 2016 were new to the project and had/did not participate in the PLD workshops.

Numbers of Cohort 2 Students

The 2016 Cohort 2 Intervention group began with 127 students in February that year. The Intervention^{+} group commenced the year with 150 students. Students in the Comparison group numbered 65 at the start of 2016. Not all students had complete data at the start of the year.

Cohort 2 Sample Characteristics

Gender characteristics of the samples are as follows:

- Intervention: 55% males; 45% females
- Intervention^{+}: 47% males; 53% females
- Comparison: 59% males; 41% females

School decile band characteristics for each sample in Cohort 2 at the start of 2016 were as follows:

- Intervention: Low = 35%; mid = 36%; high = 29%
- Intervention^{+}: Low = 23%; mid = 70%; high = 7%
- Comparison: Low = 26%; mid = 34%; high = 40%

In terms of ethnic background, the following percentages were in each of the samples:

- Intervention: Pākehā = 61%; Māori = 27%; Pasifika = 5%; Other = 7%
- Intervention^{+}: Pākehā = 63%; Māori = 28%; Pasifika = 5%; Other = 4%
- Comparison: Pākehā = 53%; Māori = 27%; Pasifika = 8%; Other = 11%

WHAT ASSESSMENTS DID WE USE AT THE START OF YEAR 1 FOR COHORT 2?

The assessments we used at the start of Year 1 were similar to those used with Cohort 1 in February 2015.

The assessments included the following:

- Letter identification
- Receptive vocabulary knowledge (BPVS)
- Word recognition (Clay Word Test)
- Invented spelling
- Phonological awareness (CTOPP)

Descriptions of these assessments are presented in Appendix 1.

What Were the Results of Assessments at the Start of Year 1?

We analysed school entry assessments by means of a series of MANOVAs for the following clusters of variables:

- Letter identification: letter names, letter sounds, upper & lower case
- Phonological awareness: elision, matching, blending
- Reading & spelling: Clay word test, invented spelling

We also analysed the BPVS standardised scores using a univariate analysis of variance.

Results of the analyses revealed that the three groups (Intervention, Intervention^{+}, Comparison) were functionally equivalent on school entry. None of the main effects for Group in the MANOVAs was statistically significant.

For receptive vocabulary (BPVS), differences in means between the three groups were not statistically significant.

Means and standard deviations for entry assessments are presented in Appendix 13a and technical data from the analyses of variance are presented in Appendix 13b.

COHORT 2 END OF YEAR 1 DATA

Sample Characteristics

Although there were 342 students across the three groups at the start of Year 1, data were obtained from 310 students. At the end of Year 1, 298 students participated in most (but not all) of the assessments. This represents an effective attrition rate of 4%.

School decile band characteristics for each sample in this cohort were as follows:

- Intervention = 109: low decile = 36%; middle decile = 35%; high decile = 29%
- Intervention^{+} = 132: low decile = 25%; middle decile = 67%; high decile = 8%
- Comparison = 60: low decile = 25%; middle decile = 35%; high decile = 40%

We include information on gender and also ethnic background. Regarding the latter, there were too few students from Pasifika, Asian and "other" backgrounds to include in our analyses; accordingly, these numbers are not included here.

- Intervention = 109 (males = 57%, females = 43%; Pākehā = 63%; Māori = 23%)
- Intervention^{+} = 132 (males = 45%, females = 55%; Pākehā = 66%; Māori = 27%)
- Comparison = 60 (males = 63%, females = 37%; Pākehā a = 54%; Māori = 26%)

Not all students had complete assessment or background data.

What Assessments Were Used with Cohort 2 Students at the end of Year 1?

The following clusters of assessments were conducted at the end of Year 1. These were the same assessments used with Cohort 1. They are described in Appendix 5.

Phonological Awareness

- Elision
- Blending
- Matching

Alphabetic Coding

- Blends
- Digraphs
- Spelling phonemes

Language Processing

- Non-word repetition task
- Mispronunciation task

Reading and Spelling

- Burt word test
- Pseudoword reading test
- Spelling

Reading Book Level

WHAT WERE THE RESULTS FOR THE COHORT 2 END OF YEAR 1 ASSESSMENTS?

We used the MANOVA Group by Decile Band design adopted in earlier analyses presented in this report.

The following clusters of variables, previously described, were treated with the MANOVAs: phonological awareness, alphabetic coding, language processing, and reading & spelling. We used a univariate analysis of variance to treat Reading Book Level data.

Phonological Awareness

The Intervention group obtained significantly higher scores than the Comparison group on each of the three phonological awareness assessments. The Intervention^{+} group outperformed the Comparison group only on the blending variable.

Regarding Decile Band, low decile students performed significantly less well overall than middle and high decile students on all three variables.

Further, and importantly, the low decile Intervention group significantly outperformed the low decile Comparison group on each of the three phonological awareness variables.

Figure 3 illustrates the results for the three phonological awareness variables, and shows how the low decile Intervention group performed at levels similar to middle and high decile students in the Intervention^{+} and Comparison groups.

Means and standard deviations for the phonological awareness variables are presented in Appendix 14a, and analyses of variance data are presented in Appendix 14b.

Alphabetic Coding

In regards to Alphabetic Coding, the Intervention group obtained higher scores than the Comparison group for each of the three assessments within this cluster. Intervention students also obtained higher scores than Intervention^{+} students on the Blends assessment.

Low decile students obtained lower scores than middle and high decile students on all three variables.

As with the Phonological Awareness variables, the low decile Intervention group obtained significantly higher scores than the low decile comparison group on each of the three variables than the Alphabetic Coding cluster.

Figures 3 and 4 on the following page illustrate group differences for Alphabetic Coding. Means and standard deviations for these are presented in Appendix 14a, and analyses of variance data are presented in Appendix 14b.

Figure 3 Year 1 Group by Decile Band Comparisons for Phonological Awareness Variables

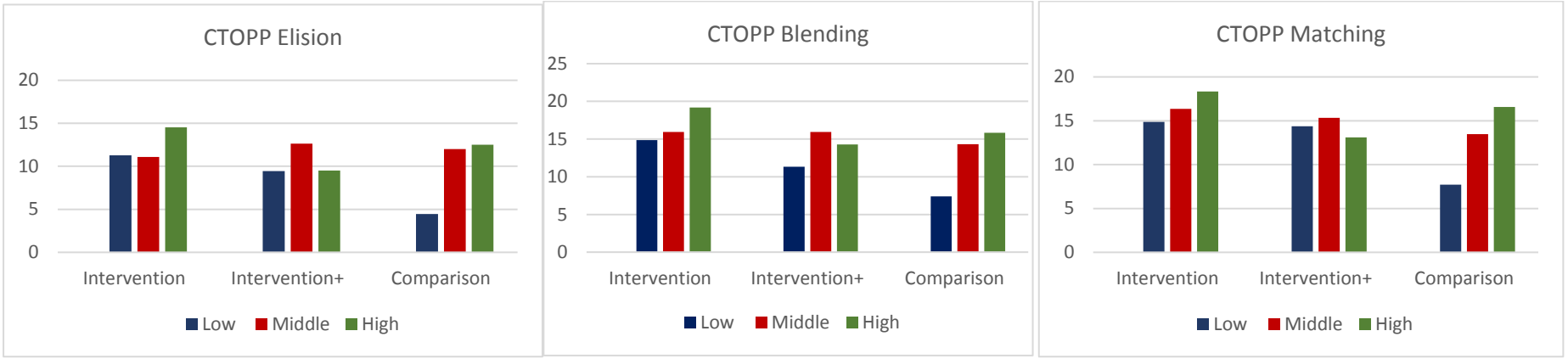
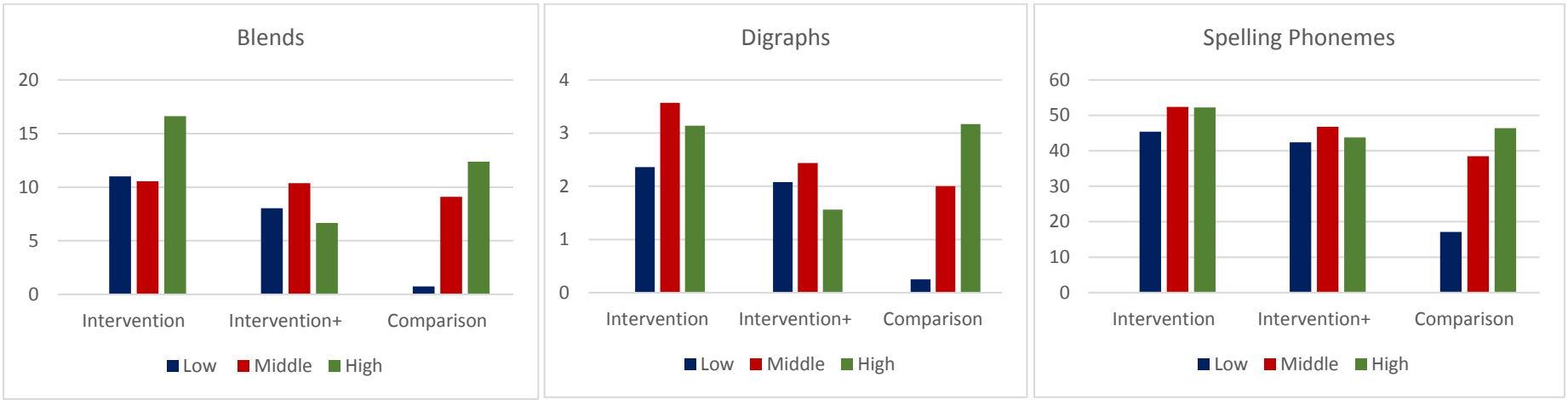


Figure 4. End Year 1 Group by Decile Band Comparisons for Alphabetic Coding Variables



Language Processing

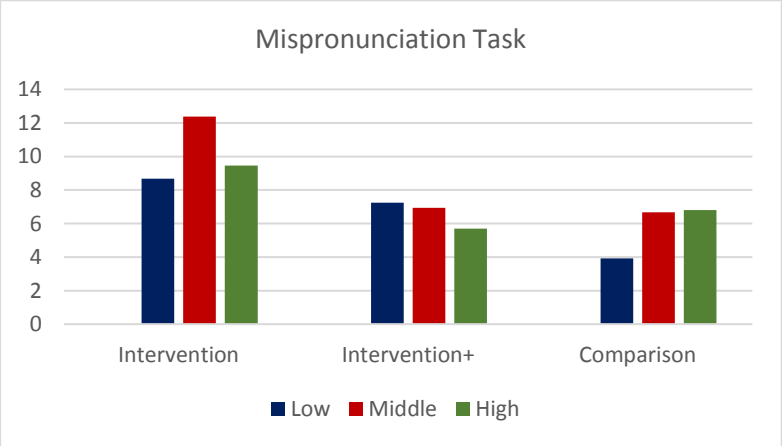
An overall significant group difference was found for Language Processing. This was due to the Intervention group obtaining significantly higher scores than the Intervention+ and Comparison groups on the Mispronunciation task.

The multivariate analysis did not result in a significant difference between the groups on the Nonword repetition task. However, there was a statistically significant effect for the Group by Decile Band interaction analysis. This result was due to two factors:

- the low decile Intervention and Intervention+ groups obtained significantly higher scores than the low decile Comparison group;
- the high decile Intervention and Comparison groups obtained higher scores for Nonword repetition than the Intervention+ group.

The group differences for the Mispronunciation task are illustrated in Figure 5. Means and standard deviations for the Language Processing variables are presented in Appendix 14a, and analyses of variance technical data are presented in Appendix 14b.

Figure 5. End Year 1 Group by Decile Band Comparisons for Mispronunciation Task



Reading and Spelling

The Intervention group obtained significantly higher scores than the Comparison group on all three variables in the Reading and Spelling cluster, and higher scores than the Intervention+ group for Pseudoword Reading and Spelling.

Further, the low decile Intervention students obtained significantly higher scores for the Burt word test than the low decile Comparison students.

Means and standard deviations for Reading and Spelling variables are included in Appendix 14a, and technical results from the analyses of variance are presented in Appendix 14b.

These results are illustrated in the following figures.

Figure 6. End Year 1 Group by Decile Band Comparisons for Burt Word Test

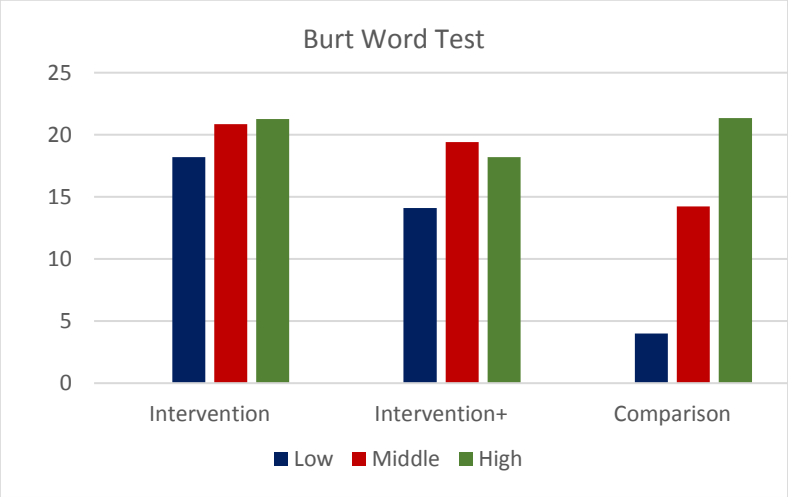


Figure 7. End Year 1 Group by Decile Band Comparison for Pseudoword Reading

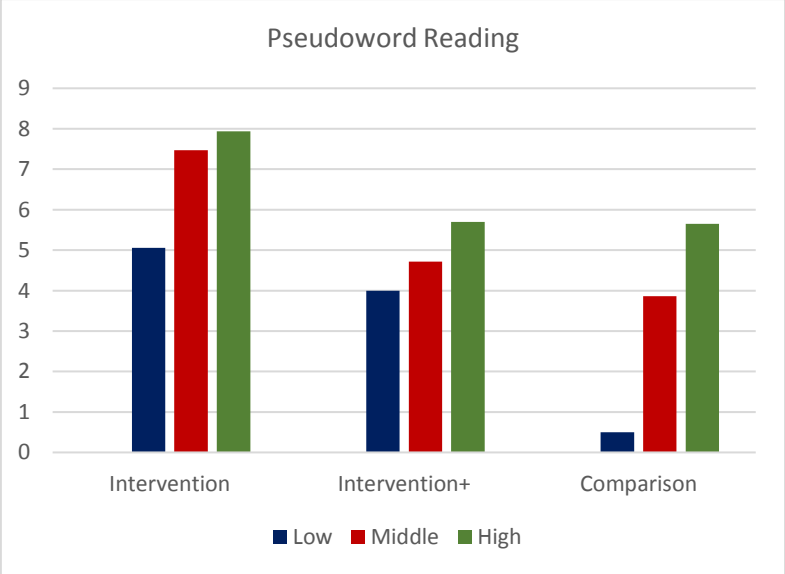
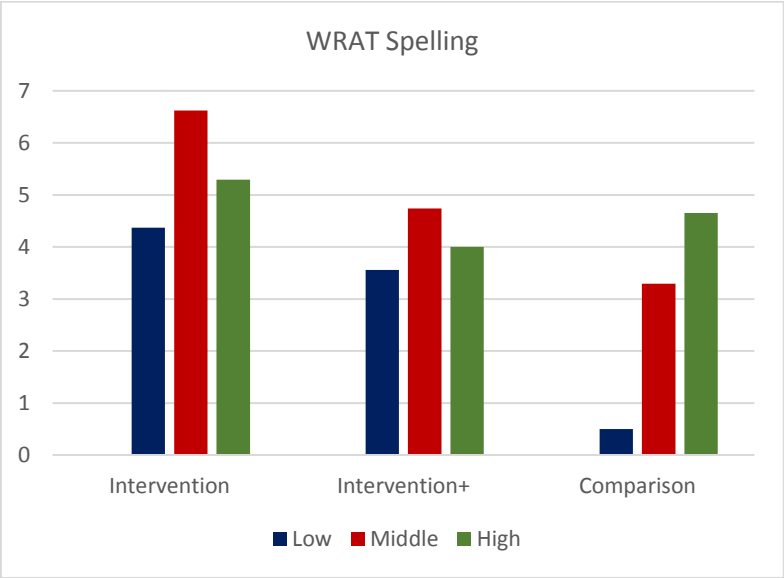


Figure 8. End Year 1 Group by Decile Band Comparisons for WRAT Spelling



Regarding reading book level, a number of schools did not provide book level data for the project, despite repeated requests. This resulted in 87 students at the end of Year 1 having no book level recorded for them. Consequently, some cell sizes were very low and including this variable in the Group by Decile Band multivariate analysis of variance design was inappropriate. Instead, we performed a one-way ANOVA to determine whether there were differences in book levels between the three groups. There were no statistically significant differences between the three groups for this variable.

Summary Cohort 2 End of Year 1 Results

Overall, the results show that the Intervention group outperformed the Comparison group on almost all variables that were assessed at the end of Year 1. Differences in favour of the Intervention group were found for the foundational skills clustered as phonological awareness, alphabetic coding, and language processing. In turn, these skills were reflected in the superior performance of the Intervention group on the word reading and spelling assessments.

Particularly noteworthy are the results for the low decile Intervention students. On a number of assessments their performance approached that of students in higher deciles. The low decile Intervention students also dramatically outperformed the low decile

Comparison students on virtually all assessments. Considering that the low decile Comparison students generally performed at similar levels to low decile Intervention students at the start of Year 1, the result is impressive for these low decile Intervention students.

COHORT 2 MIDDLE OF YEAR 2 DATA

Sample Characteristics

At the final data collection point in this project, the middle of Year 2 (2017), 281 students were still in schools that were participating in research. The overall attrition rate of Cohort 2 students in terms of those who first participated in the beginning of Year 1 assessments (February 2016) was around 9%.

School decile band characteristics for each sample in this cohort were as follows:

- Intervention = 98: low decile = 35%; middle decile = 35%; high decile = 30%
- Intervention^{+} = 130: low decile = 25%; middle decile = 68%; high decile = 7%
- Comparison = 53: low decile = 23%; middle decile = 32%; high decile = 45%

We include information on gender and also ethnic background. Regarding the latter, there were too few students from Pasifika, Asian and "other" backgrounds to include in our analyses; accordingly, these numbers are not included here. Further, schools were unable to provide background data for a number of students.

- Intervention = 98 (males = 56%, females = 44%; Pākehā = 64%; Māori = 22%)
- Intervention^{+} =130 (males = 46%, females = 54%; Pākehā = 66%; Māori = 27%)
- Comparison = 53 (males = 59%, females = 41%; Pākehā = 57%; Māori = 24%)

What Assessments Were Used with Cohort 2 Students at the Middle of Year 2?

The following clusters of assessments were conducted at the middle of Year 2. The location of descriptions for each assessment is shown in parentheses.

Language comprehension

- Listening comprehension (Appendix 9)
- Nonword repetition (Appendix 5)
- Morpheme awareness (Appendix 11)
- Receptive vocabulary (BPVS: Appendix 1)

Reading and Spelling

- Burt word test (Appendix 5)
- Pseudoword reading test (Appendix 5)
- Wide Range Achievement Test (WRAT): Spelling (Appendix 9)

Reading Book Level (Appendix 5)

Word Attack Strategy (Appendix 11)

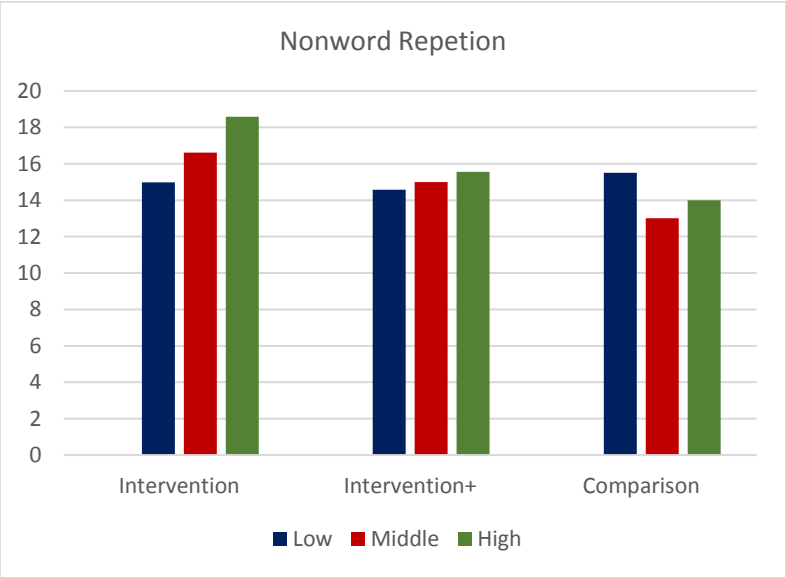
Reading Self-Efficacy (Appendix 11)

What Were the Cohort 2 Results at the Middle of Year 2?

Language Comprehension

There was a marginally significant group difference (p = .06) in the multivariate analysis of this cluster of variables. The difference was due to the Intervention group obtaining significantly higher scores than the Intervention^{+} and Comparison groups for the Nonword repetition variable. This finding is illustrated in Figure 9.

Figure 9. Middle Year 2 Group by Decile Band Comparison for Nonword Repetition



Although the Intervention group obtained higher scores than the Intervention+ and the Comparison groups for the other variables in this cluster, the differences were not statistically significant.

In line with findings throughout the project, students in low decile schools performed less well than students in middle and high decile schools on all but one variable in this cluster. The exception was for Nonword Repetition, in which the differences between the Decile bands were not statistically significant.

Means and standard deviations for Middle of Year 2 variables are presented in Appendix 15a. Technical data from the analyses of variance are presented in Appendix 15b.

Word Recognition

There was a strong, overall group difference for the cluster of word recognition variables. The Intervention group obtained significantly higher scores than the Comparison group on each of the three variables, and higher than the Intervention+ group for Pseudoword reading. This is an important result.

Similarly, there were significant differences between the Decile band groups on all variables. Low decile students obtained lower scores than middle and high decile students.

Although the low decile Intervention students obtained higher scores than the low decile Comparison students, these differences were not statistically significant. The following three figures illustrate the group by decile band differences for each variable.

Figure 10. Middle Year 2 Group by Decile Band Comparisons for Burt Word Test

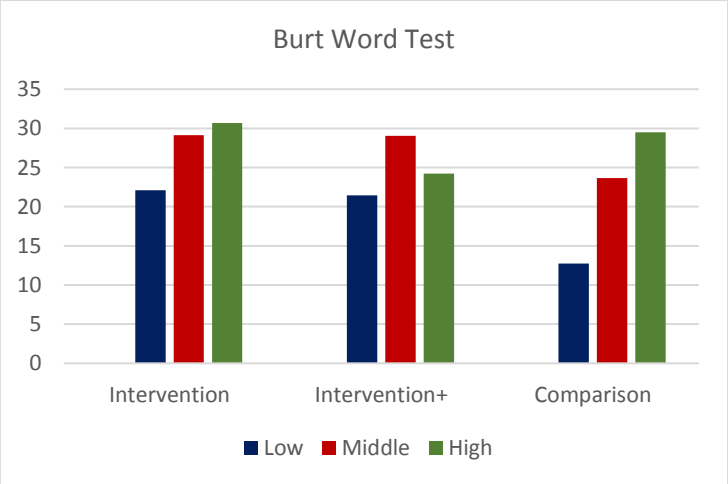


Figure 11. Middle Year 2 Group by Decile Band Comparisons for Pseudoword Reading

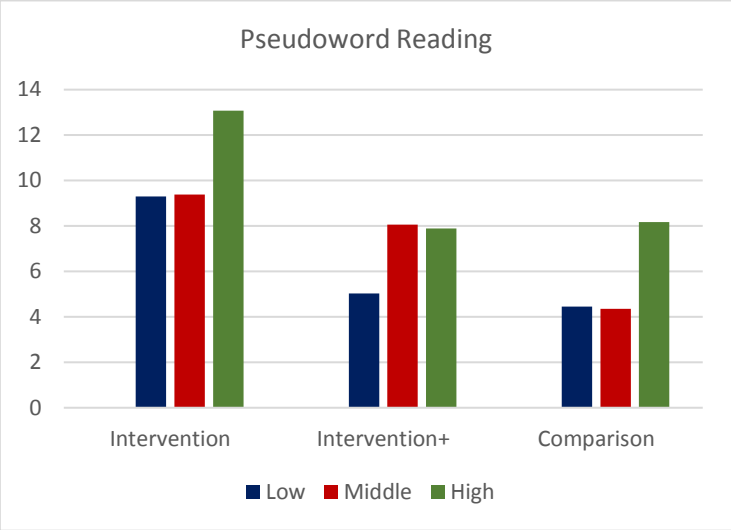
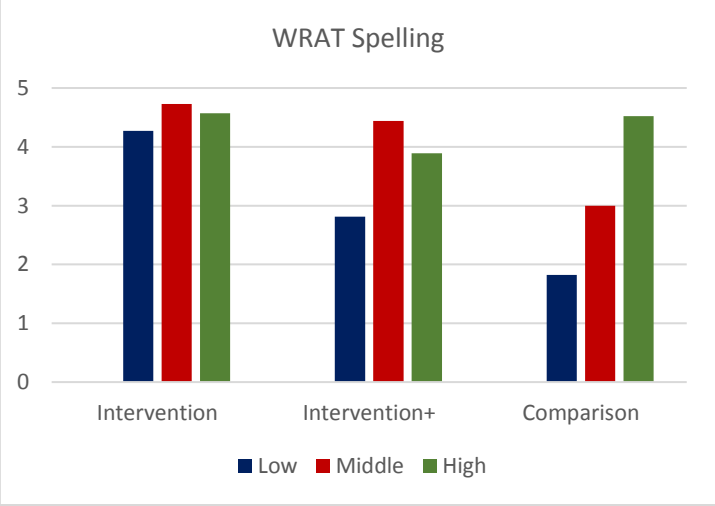


Figure 12. Middle Year 2 Group by Decile Band Comparisons for WRAT Spelling



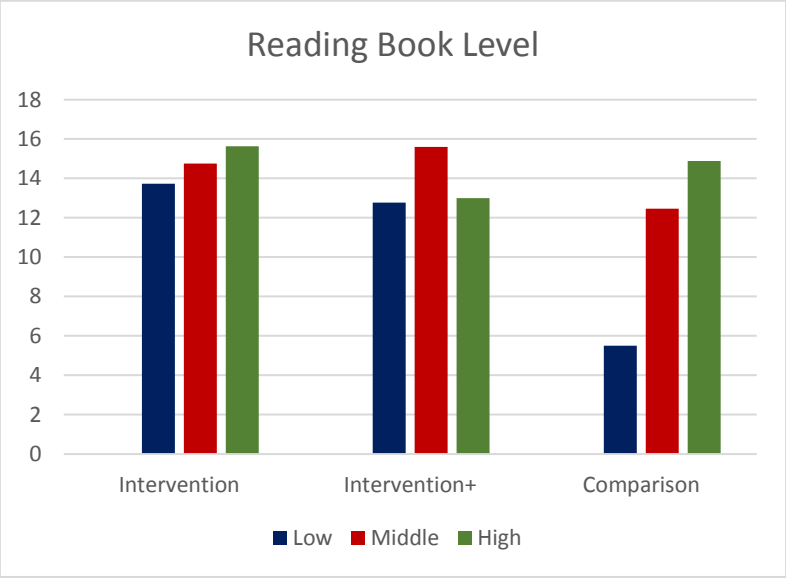
Reading Book Level

Schools supplied Reading Book Level data for 260 of the 281 students who remained in the project at the middle of Year 2. The analyses reveal that the Intervention group obtained significantly higher book levels than the Comparison group. The low decile Intervention students significantly outperformed the low decile Comparison students. The group differences are illustrated in Figure 13.

Across the three groups combined, low decile students obtained lower book levels than middle and high decile students, due largely to the results for the low decile Comparison students.

Means and standard deviations for book levels are presented in Appendix 15a, and technical results from the analyses of variance are presented in Appendix 15b.

Figure 13. Middle Year 2 Group by Decile Band Comparisons for Reading Book Level



Word Attack Strategies

The word attack task asked students to tell what they do when they are reading on their own and come across a word they don't know. Responses are scored as either being a word-specific strategy (sounding out; looking at the letters) or a non-word-specific strategy (asking someone, guessing, looking at the picture).

Responses were available for 219 students, and were analysed by means of a Chi-square. Overall, 83% of responses were word-specific. The Chi-square analysis indicated that there were no statistically significant differences between the groups.

Reading Self-Efficacy

There were no statistically significant differences between the groups or between the decile bands for the Reading Self-Efficacy Scale. The Intervention group was higher than the Intervention^{+} and Comparison groups, but the differences were small and non-

significant. Data are presented in Appendix 15a, with analysis of variance results appearing in Appendix 15b.

Were There Differences Between the Two Intervention Cohorts?

We conducted analyses to determine whether there were differences between the two Intervention cohorts on reading and spelling assessments at a similar time, namely the middle of Year 2. This assessment point was the last stage of data collection for Cohort 2.

The multivariate analysis of the three reading and spelling variables revealed a statistically significant difference between the two Intervention cohorts. These differences were significant for pseudoword reading and WRAT spelling. In both cases, the Cohort 2 Intervention group obtained significantly higher scores than the Cohort 1 Intervention group. Differences for the Burt word test showed higher scores for the Cohort 2 Intervention group, but these were not statistically significant.

The following figures illustrate the findings from these Cohort Intervention group comparisons. Means and standard deviations and technical analyses of variance data are presented in Appendices 16a and 16b respectively.

Figure 14. Middle Year 2 Intervention Cohorts by Decile Band Comparisons for Burt Word Test

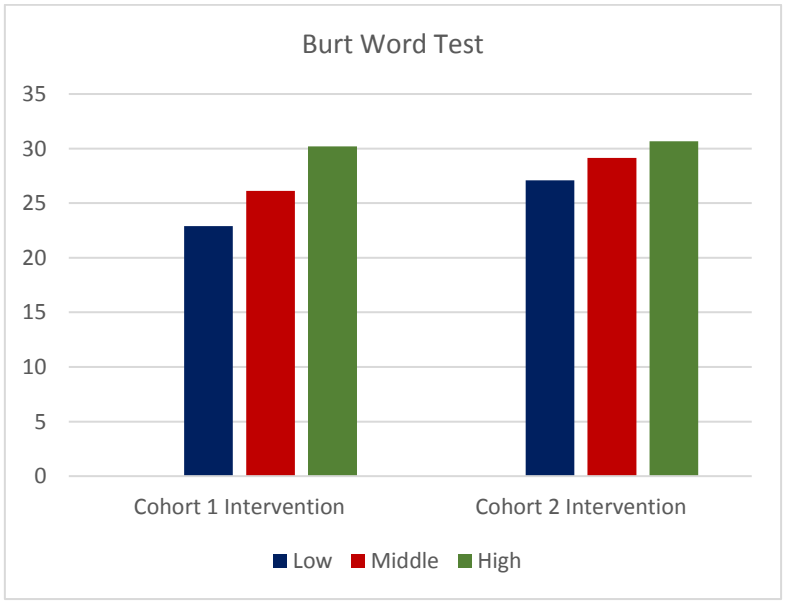


Figure 15. Middle Year 2 Intervention Cohorts by Decile Band Comparisons for Pseudoword Reading.

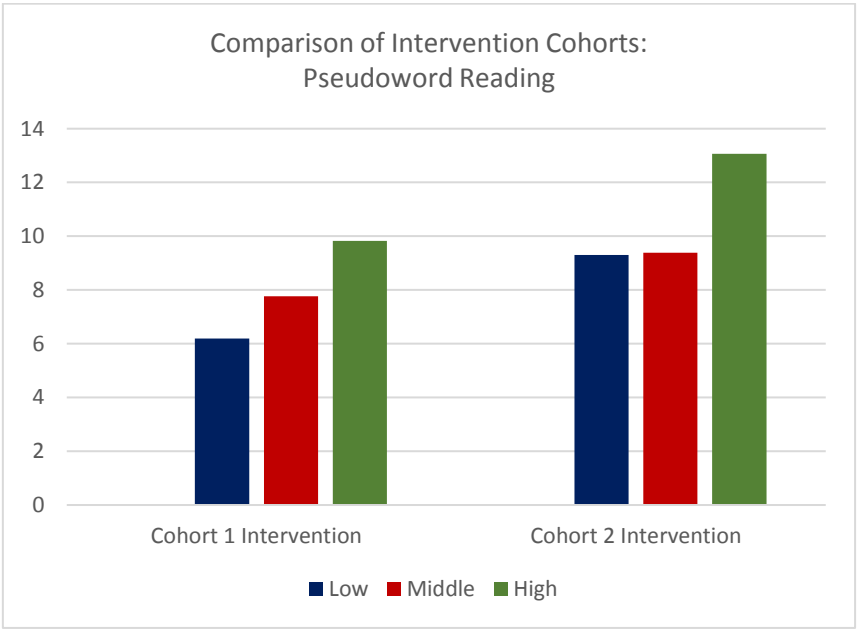
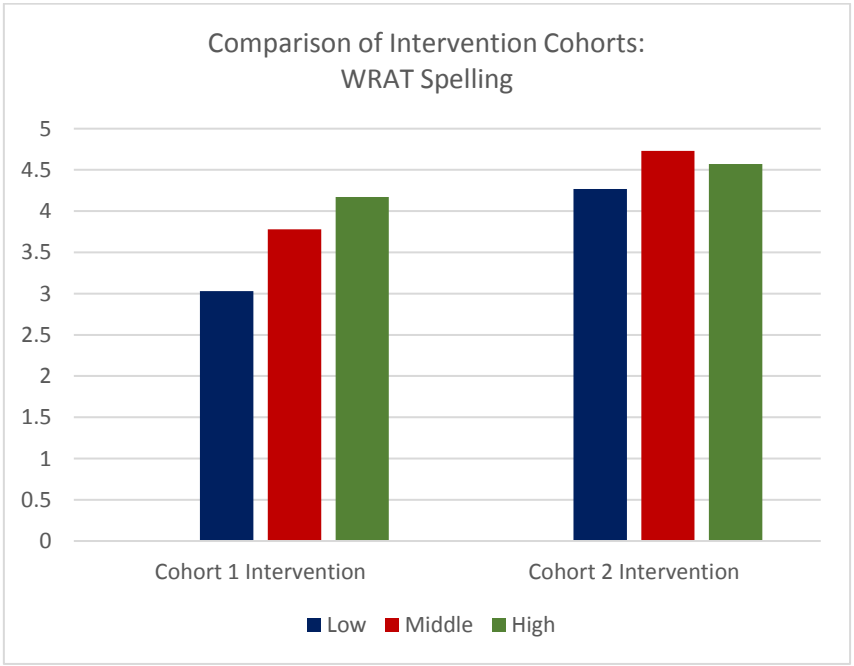


Figure 16. Middle Year 2 Intervention Cohorts by Decile Band Comparisons for WRAT Spelling.



SUMMARY OF COHORT 2 RESULTS

The Cohort 2 results indicate the following key findings:

1. The Intervention, Intervention^{+} and Comparison groups were functionally equivalent on school entry at the start of Year 1. On a number of variables, the Comparison group had slightly higher scores than the Intervention groups, but these were not statistically significant.
2. At the end of Year 1, clear differences emerged. The Intervention group outperformed the Comparison group on almost all measures, and were better than the Intervention^{+} group on many variables. Of particular significance is the finding that for the reading "outcome" variables (Burt word test, pseudoword reading, spelling), the Intervention students were superior to the Comparison students. Especially notable was the strong performance of the low decile Intervention students.
3. At the middle of Year 2, marking the last data collection point for Cohort 2, the Intervention students outperformed the Comparison students on the three Reading and Spelling assessments and one aspect of Language processing (nonword repetition). Again, the low decile Intervention group was superior to the low decile Comparison group. For Reading Book Level, differences were not statistically significant, although the low decile Intervention students obtained higher book levels than the low decile Comparison students.
4. There were no statistically significant differences between the groups for Reading Self-Efficacy. Scores on this measure were generally high (positive). It is perplexing that the highest score was obtained by the low decile Comparison group, which consistently had the lowest scores on almost all variables.
5. Our expectation that Cohort 1 Intervention teachers might consolidate PLD learning and achieve better results with another group of New Entrant/Year 1 students (Intervention^{+} group) was not fulfilled. The Intervention^{+} students did not achieve significantly better results than the Comparison students on almost all variables.

6. Most variables at all assessment points resulted in overall differences (collapsed across groups) in scores for Decile Band groupings: low decile students typically obtained lower scores than high, and sometimes middle, decile students.
7. Comparing middle of Year 2 means for the two Intervention cohorts on the reading and spelling assessments revealed that the Cohort 2 Intervention group outperformed the Cohort 1 Intervention group.
8. Changes made to the format and delivery of PLD workshops with Cohort 2 teachers appear to have been successful and are associated with significantly improved literacy learning outcomes for Cohort 2 students compared to Comparison and Cohort 1 Intervention students.

REPORT ON TEACHER DATA FOR COHORT 2

By Christine Braid

The clear differences in student outcomes between the Cohort 2 Intervention and Comparison groups are associated with Intervention teacher data changes from pre-intervention to post-intervention.

As with Cohort 1, teacher data were gathered on teacher knowledge and teaching practice. Teacher knowledge was measured using the Basic Linguistic Constructs test and a self-evaluation survey as used and described for Cohort 1 (Appendix 2). Data on teaching practice were gathered from video of small group reading instruction and teacher word identification prompts scenarios. The videos of teaching were analysed by a rubric developed for the project. The rubric is described in Appendix 17. The prompts scenario, described in Appendix 17, was adapted from the Greaney (2001) version used with Cohort 1.

Teacher Knowledge

We carried out a survey on two aspects of teacher knowledge: their linguistic knowledge and their self-evaluations of knowledge for literacy teaching. Linguistic knowledge was tested using a Basic Linguistic Constructs test and self-evaluation data were gathered using a Likert scale.

Basic Linguistic Construct Test

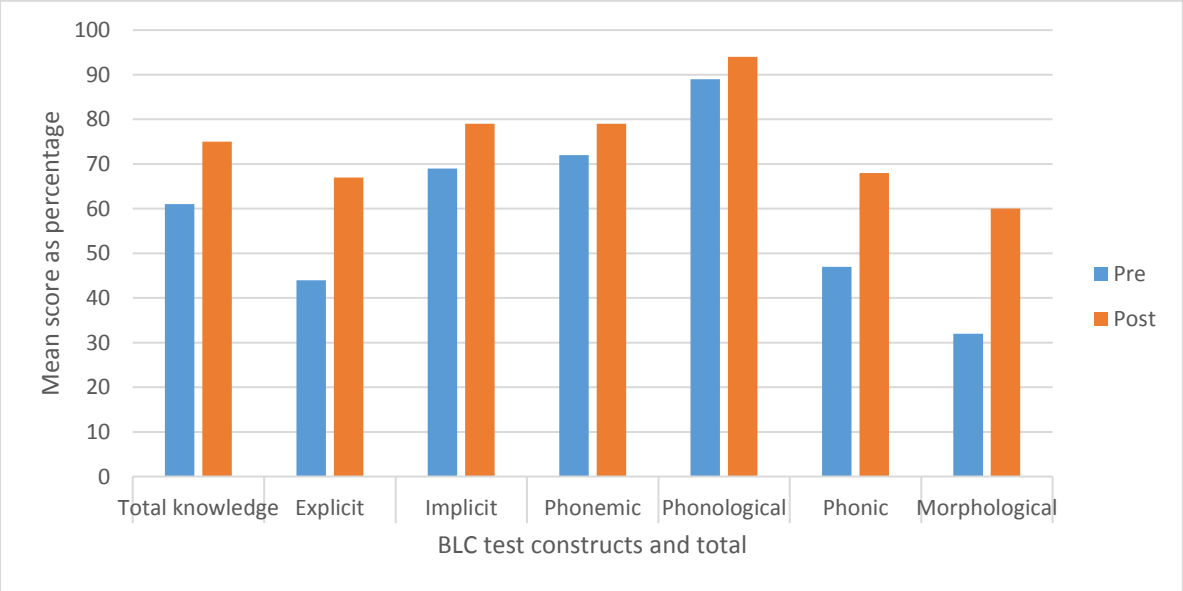
The Basic Linguistic Constructs test, as described in Appendix 2, was used with intervention teachers in Cohort 2 to identify levels of teacher knowledge in constructs determined important for effective teaching of beginning reading. We administered the test in the first workshop to provide pre-intervention data, and in the final workshop to provide post-intervention data. Teachers were asked to complete the test individually and to answer all questions even if their response was ‘I don’t know’. As outlined in Appendix 2, the test covered four key linguistic constructs: phonemic, phonologic, phonics, and morphologic.

Teachers' phonemic and phonological knowledge are important for understanding how the sounds of language map onto the written code (phonics). Phonic and morphological knowledge enable teachers to understand the complexities of English orthography and how to best teach this to their learners. Implicit and explicit knowledge are necessary for teachers to identify student need and create appropriate teaching opportunities.

Results were available from 20 teachers who had pre- and post-intervention test data. At Time 1, teachers scored an average of 61% correct across 38 items tested, a score similar to that of Cohort 1 teachers' pre-intervention scores. As with Cohort 1, teachers' implicit knowledge was higher than explicit, and phonemic and phonological items were high in comparison to phonic and morphological knowledge. Teachers' implicit knowledge was higher than explicit knowledge, and phonemic and phonological items were high in comparison to phonic and morphological knowledge.

In the post-intervention test, Cohort 2 teachers' scores had increased to an average of 75% correct. This improvement was considerably greater than the performance of the Cohort 1 intervention teachers, whose results were similar to the Cohort 2 teachers' results at Time 1, and remained similar at Time 2. Figure 17 illustrates the Time 1 and Time 2 results for Cohort 2 teachers.

Figure 17 **Basic Linguistic Constructs test results for Percentage of Items Correct at Time 1 and Time 2**



Paired samples *t*-tests showed the increases in average total scores were statistically significant, t (19) = 5.35, $p < .01$, d (effect size) = 1.00. All t test results for each dimension of the BLC showed a statistically significant increase at Time 2 compared to Time 1. The results are presented in Appendix 18.

Teachers had high levels of knowledge of phonemic constructs (>70%) at both testing periods, with most teachers able to correctly identify how many phonemes are in a word, and define a *phoneme* as the smallest speech unit in a word. Some confusion existed in specific concepts such as the relationship between the speech sounds involved in consonant blends, digraphs, and the letter 'x'. Confusion with the phonemes in the letter 'x' (which has a [ks] sound) was also seen in teacher video where a teacher struggled to find an example that helped students with the phonemic interpretation of the grapheme 'x'.

Phonological knowledge was also high (>85%) at both time periods, with teachers successful in identifying how many syllables in a word. Teachers scored lower in identifying syllable types, which showed that explicit knowledge was low in comparison to implicit knowledge.

Another explicit knowledge item that teachers had difficulty with at Time 1 was in selecting a correct definition for phonemic and phonological awareness. Teachers selected the definition that phonemic awareness was "*understanding how letters and sounds are put together to form words"* and the definition of "*ability to use letter sound correspondences to decode"* for phonological awareness. Both answers show teachers' confusion between phonological and phonics concepts. At time 2, the confusion remained for defining phonemic awareness but responses increased to an average of 65% correct for defining phonological awareness.

Initial high levels of phonemic and phonological knowledge are evidence that for most teachers professional learning has been successful in these aspects. As we have noted elsewhere in this report, many teachers use a programme for teaching phonics (e.g., Yolanda Soryl, Jolly Phonics) and along with training in Elkonin boxes (for example in Reading Recovery) and hearing sounds in words when writing, this may have helped teachers develop knowledge in these foundations of language.

Teachers' phonic knowledge scores averaged 47% correct at Time 1, similar to the results in Cohort 1. The average scores increased to 68% at Time 2, showing that teacher

knowledge had increased but some gaps in phonic knowledge remained. The answers to test items revealed that teachers were confused about the difference between consonant blends and consonant digraphs. This confusion was also evident in the phonemic items of the test and in some teacher video, where, for example, teachers and students co-constructed a list of consonant blends that included some consonant digraphs.

Teachers' knowledge of morphological constructs was the lowest of average scores; 32% at Time 1, but increased significantly to 60% at time 2. Teachers' ability to define a morpheme increased from an average of 40% at time 1 to 80% at Time 2, and correct responses for counting morphemes increased from an average of 35% to 75%. The low levels of phonic and morphological knowledge at Time 1 likely constrain teachers' ability to effectively teach code.

The results indicate that the PLD workshops clearly had a positive impact on teacher knowledge. Notwithstanding that some gaps in knowledge remained at Time 2, the increase in scores is important and meaningful.

It was interesting to note that many teachers' answers to one question revealed a belief that spelling rules were random rather than governed by rules of patterns. Teachers' comments in workshops indicated that it was common for teachers to believe that English is unreliable in its spellings. Such a belief will mean that teachers are unlikely to understand that English orthography can be taught in a systematic and sequenced way. In the workshops, teachers were presented with facts that English spellings are 97% regular when orthographic patterns and morphology are considered alongside phonology (Crystal, 2000).

Teacher Self Evaluation

To complement the teacher knowledge survey, teachers were asked to evaluate their ability on eight aspects of teaching literacy. The first five aspects were phonological, phonics, vocabulary, fluency, and comprehension. These five aspects were identified by the National Reading Panel (2000) as vital elements in literacy teaching and learning. In addition, teachers evaluated three classroom variables for using reading assessments, children's literature, and for teaching English Language Learners. Teachers rated their ability as very low (1), low (2), good (3), or very good (4).

At Time 1, teachers rated their ability to teach comprehension higher than the teaching phonemic and phonics knowledge. At time 2, the statistically significant increases

in self-evaluation of knowledge in phonemics and phonics, and the lower level of increase for comprehension resulted in a closer ranking between the code and comprehension elements. The results show that at Time 2, teachers reported considerably greater confidence in the code elements of reading than they had at Time 1. Appendix 19 presents a table of these self-evaluation results.

The changes in teacher self-evaluation are important for a number of reasons. Both code and comprehension components are vital for successful reading (Tunmer & Hoover, 2014). Students' early success with code knowledge is particularly important considering the role this plays in independence and on-going self-teaching about words (Share, 1995).

Teaching practice

We analysed observations of teaching practice in small group reading instruction, and the approach teachers reported using in helping a child in the process of reading (teacher prompts). An observation rubric was developed to facilitate analysis of small group teaching practice. A prompts scenario was developed to analyse teachers' approach to assisting students to correct mistakes or identify unfamiliar words during reading. Both the observation and the prompts tools, as outlined in Appendix 17, position an explicit focus on code knowledge as important.

Observations

In small group reading instruction, the focus of the intervention workshops was on an explicit and systematic approach to word learning, as opposed to the common practice of guided reading with a focus on using a levelled book. We used the rubric (Appendix 20) to categorise and analyse the teaching practice at Time 1 and Time 2. The rubric in Table 3 presents an overview of the categories.

Table 3: *Overview of the key differences between implicit and explicit lessons from the observation scale*

	Incidental ----------Implicit 1 2	Explicit ----------Systematic 3 4
Lesson focus	Making meaning from instructional text	Explicit teaching of knowledge and strategy to apply to text
Instructional strategies	Questioning ------prompting	Modelling----explaining
Code knowledge	As arising in instructional text	Explicit ------Scope and sequence
Text selection	Natural ------- Controlled	Specific -------decodable
Reader strategies	Meaning cue and first letter	Code cue first ----decode strategy
Materials used	Instructional text word cards	Whiteboard; magnetic letters

Observation data were available from 21 teachers who had video of practice prior to and after the PLD intervention programme. The pre-intervention teaching practice was more closely aligned with a guided reading and whole language approach than with a cognitive information processing approach. The whole language approach involves the use of a connected text as dominant in teaching a child to read and any teaching of words occurs during the reading of meaningful text (Clay, 2005). Such practice aligns with a categorisation of 1 and 2 on the rubric.

A cognitive information processing approach involves the explicit teaching of skills required for successful reading. Such an approach was promoted by the intervention workshops. The reading of a text provides an opportunity to practice the skills taught. Such teaching practice aligns with a categorisation of 3 and 4 on the rubric.

The results from the observation rubric showed that prior to the intervention, teaching practice in small group instruction followed an implicit approach, categorised as incidental (1) or implicit (2), in 100% of the cases. No lessons were categorised overall on the explicit end of the rubric. The results for pre-intervention are expected because the

implicit approach is the dominant discourse about the teaching of reading in New Zealand, as seen in teacher materials and in school practices.

The results at Time 2 showed a meaningful change to teaching practice that was categorised at the explicit end of the observation continuum, in either category 3 or 4. One lesson (5%) was categorised as systematic (category 4) and 43% of lessons as explicit (category 3). However, it is worth noting that 52% of lessons remained at the implicit end of the rubric in either category 1 (24%) or category 2 (28%). Many teachers had difficulty with some of the changes suggested to small group reading instruction. Appendix 21 presents the results for changes in the lessons for each element of the rubric.

The changes to teaching practice that teachers made most easily were in the elements of lesson focus and materials used. Explicit and systematic teaching was seen in 71% of lessons for lesson focus, and 67% of lessons for materials post-intervention. Teachers were able to modify their lessons by adding a focused teaching time that used materials such as magnetic letters. Text selection and teaching strategies were the elements that remained more implicit in approach, with 52% of text selection and 43% of teaching strategies remaining implicit or incidental.

The results show that teachers appeared to have difficulty in finding suitable texts to support the explicit teaching and application of effective decoding strategies. It was apparent that teachers had embedded practices involving guided reading and the associated commonly used texts. In teaching students strategy use, changing to explanation and modelling instead of a predominance of teacher questions was difficult for many teachers.

Examples of incidental (1) and implicit (2) practice from the video observations showed that teachers often introduced the lesson with an introduction to the levelled book to be read. In some cases, a learning intention for how to apply the reading process was displayed and discussed prior to the reading. For example, one learning intention was to “think about the story and use the pictures and the first letter of the word”. Learning intentions were in such cases used as the specific teaching prior to introducing a text.

In the incidental and implicit lessons, word teaching occurred in the context of story reading, which resulted in large amounts of incidental teaching in order to support students to successfully read the text. As examples of such teaching, a sentence of 11 words required teacher support for eight of those words, in a level 6 text. In another sentence, one

particular word was focused on for 70 seconds as the teacher assisted the students to decode it. In a lesson using a level 14 text, one teacher assisted the students with eight of the 23 words in a sentence. The eight words covered six different orthographic patterns and the sentence took just under three minutes to process.

Other practices common to the implicit lessons were teachers reading each page of the book to the child as a detailed introduction to the text and many examples of teacher sub-articulation of a word the students had stopped on.

Examples of lessons rated as explicit (3) or systematic (4) showed that teachers used the first part of a lesson for explicit instruction of knowledge and/or strategy appropriate to the students' phase. The section involved teacher explanation and modelling. For example, *"This is the letter D; it makes the sound /d/".* The teachers who used this approach then reduced teacher input gradually over the lesson moving on to questions *"What letter is this? What sound does it make?"*

Other examples of an explicit teaching section involved the teachers showing strategies for the students to learn, explaining and modelling how a consonant-vowel-consonant (c-v-c) combination can be sounded and blended to make a word. Teachers used magnetic letters to model sounding and blending and allowed students to engage with the task and materials. Lessons in these categories utilized the teaching section for practice and over-learning by using as much repetition as possible. A text was used to allow the students to apply the strategies and the teacher expected the students to read the text independently.

The incidental approach, as described above, serves to make a child more dependent on the teacher for a successful reading of the text. In contrast, the explicit teaching scenario teaches what the child needs to successfully read the text. With practice, students will apply these types of skills and more quickly read independently.

Teacher Prompts

The results from the prompts scenario tool showed that a number of teachers made a change in practice from using meaning as the first prompt towards using word cue first. Another change was noted from teachers using prompts that were general, towards prompts that directed students to specific decoding strategies.

Teachers use of both a word cue first and specific decoding support across all five scenarios increased from 27% at time 1 to 62% at time 2. For suggesting a word cue first, the results changed from 32% of teachers at Time 1, to 52% of teachers at Time 2. For suggesting a specific decoding strategy, the results showed a change from 20% of teachers at Time 1 to 65% at Time 2.

The results indicate that teachers attempted to make changes in line with the workshop suggestions. Less change was noted in the use of a word cue first than in adding specific guidance for using decoding strategies. Asking students about meaning or sentence structure as a first cue is clearly a deeply embedded practice for most teachers. *Meaning is very elusive if the words that convey meaning are not decoded.*

Teachers' time 1 prompts involved a high proportion of questions that asked about the meaning or structure of the sentence, such as, "*Did that sound right?"* and *"Did that make sense?"* These suggested prompts offer no direction to help the reader to use a decoding strategy to solve the word; instead, they direct students to use meaning and sentence structure only. Other prompts with little or no direction towards a decoding strategy included generalised prompts such as *"Can you find your mistake?",* "*Try that again*", or "*Get your mouth ready".*

While these prompts aim to have students become independent in strategy when they make errors, they do require a high level of skill and tend to cause confusion for students during the early stages of learning to read. It could be overwhelming for students to have to find an error when print is a new concept to them. Clay (2005) suggested that teachers may use apparently good prompts at inappropriate times and outlined that the "*Get your mouth ready*" prompt "assumes the child knows where to find the first letter and has already learned the letter-sound relationship*"* (p. 93).

One problem the prompts scenario responses highlighted was that a number of teacher prompts have become embedded and almost scripted, reflecting advice and suggestions in the instructional support materials commonly used by New Zealand teachers.

Teachers need a high level of knowledge about the child's developmental phase and the specific orthographic pattern of target words to use prompts well. The evidence from the scenarios is that at Time 2, teachers were applying their new linguistic knowledge. Time 2 prompts more frequently involved teachers directing students to look at the letters first,

with associated guidance towards decoding strategies to solve the error. For example, teachers asked students to blend *o-n* to read the word 'on', or to decode *l-ea-der* and think what word would fit with the evolving meaning of the sentence.

Teacher Data and Student Outcomes

The students' results presented earlier in this report show statistically significant increases for a number of measures when comparing the Intervention and the Comparison groups. We suggest that the student improvements are associated with the changes in teacher practice and knowledge. Cohort 2 teacher knowledge at Time 1 (pre-intervention) was similar to Cohort 1 teachers. The changes in teacher knowledge and practice observed at Time 2 (post-intervention) are likely to have been the result of the intervention, resulting in a positive impact on student outcomes at the end of Year 1 and the middle of Year 2.

Interview data

To examine what teachers found to be enablers and barriers to changing practice, semi-structured interviews were carried out with four teachers. These teachers had medium to high teacher knowledge on the BLC. Teaching practice varied; one teacher had systematic practice, one had explicit practice, one had implicit practice, and one was predominantly incidental in her practice.

The interview questions were based on knowledge of the teachers' videos and helped to reveal how teachers were applying the intervention workshop information. As Table 3 shows, barriers to changing teaching practice included beliefs, set practices, student need, and time. Colleagues and resources were identified as barriers in some instances and enablers in others.

Table 4: ***Barriers to changing teaching practice***

Themes	Enabler	Barrier	Definition
Beliefs/ideology		●	Teachers' responses showed a way of thinking that is embedded and is difficult to challenge. E.g., multiple-cueing belief about the reading process
Set practices		●	Some set ways of teaching or use of programmes were resistant to challenge, e.g., guided reading.
Resources	●	●	Some resources reinforce set ways of teaching, e.g., the levelled book series. Some resources enable and support change, e.g., decodable texts and scope and sequence,
Teacher knowledge	●		Teachers' knowledge about linguistic constructs had a positive impact on explicit teaching and phases of word learning

Colleagues	●	●	Collegial support had a positive impact on teacher practice. Colleagues can also be a challenge or disallow new practice when a strong belief system cannot be changed
Student needs and progress	●	●	Positive student progress after new teaching impacts on continuing practice. When progress is attributed to teaching, it is an enabler to change. When difficulty in making progress is attributed to the child, it is a barrier.
Time		●	The time needed to embed new practice competes with many other time constraints in a teacher's day

Beliefs

It was evident that a major barrier to teachers changing their practice was the belief teachers held about reading as a multiple-cueing system, rather than as a code and comprehension driven process. The multiple cues approach has formed the basis of teachers' beliefs about how reading develops. The approach means teachers promote that students use meaning and sentence structure to enable the reading process, with word level detail used only as necessary and often as a last resort (Clay, 2005; Goodman, 2008).

Set Practices

Set teaching practices emerged as a barrier to change in practice. At Time 1, teachers favoured a guided reading approach where the book was used as the place to incidentally teach code knowledge; changing to a more explicit approach was difficult for some teachers. Many videos showed teachers added an explicit teaching section on code knowledge to their lessons and attempted to follow the scope and sequence to teach students about the code. Some teachers remained reliant on a guided reading approach and the selected levelled book for any teaching of the code.

Resources

The resources available to teachers are in many cases a barrier to changing practice. The application of a scope and sequence for word learning requires materials and texts that use a systematic approach to words used. In the early stages of reading this would mean a large number of opportunities to practice a sounding and blending of regular c-v-c words. However, the dominant book series is based on natural language, and consequently is not controlled for such opportunities. Table 4 shows how the *Ready to Read* series has fewer

opportunities to practice regular c-v-c word patterns than a similar level book from the decodable series of *Little Learners*. The table also shows that the *Ready to Read* text has difficult word patterns that the *Little Learners* series removes for early stage learning.

Table 5: ***Comparison of word patterns in two different book series***

	No. of words	c-v-c or v-c	Other short vowel	Vowel teams	Suffixes / compound	Sight words
Ready to Read	100	4	4	7	6	11
Little Learners	96	33	15	2	0	9

Enablers to changing practice

The challenges to teachers making changes to practice were in some instances offset by the supports of the intervention programme. As Table 4 shows, enablers included teacher knowledge and student progress, with resources and colleague support featuring as both a barrier and enabler.

Resources

During the workshops, teachers were provided with a scope and sequence that outlined the phases of word learning. This provided a structure for teaching. Teachers reported that they found the provision of a scope and sequence to be useful and most teachers reported using it. A common teacher comment was that the scope and sequence gave them a tool to identify and address the gaps some students had. The scope and sequence also helped teachers apply the new knowledge from workshops. New teacher knowledge was used in conjunction with the scope and sequence, enabling teachers to be more explicit in teaching about words.

Other new materials provided as support for teachers' practice included a series of decodable readers and sets of magnetic letters. Some teachers found these a support and were able to adapt lessons to use the scope and sequence and the decodable texts. Other teachers found the series difficult to use if they continued with a multiple-cueing approach to teaching reading.

Teacher Knowledge

Teachers generally engaged with the new knowledge of linguistics and the developmental nature of learning to read. The increases in teacher knowledge were evident

in teaching practice as teachers began to include explicit teaching in small group instruction. Lessons after the intervention showed teachers using their knowledge of the alphabetic principle and of decoding strategies in an explicit teaching section of small group instruction. Teachers commented that they felt the new knowledge improved their teaching. While most teachers could see the need for the high level of linguistic knowledge, some thought the knowledge was not necessary for teaching new entrant students.

Colleagues

Teachers commented that colleague support helped them to make changes to their practice. When changes were made across a team, teachers reported feeling positive about making changes and the impact they felt this would have on the students' outcomes. During the interviews, one teacher noted that she was excited about having the new knowledge and processes as something all the team would follow because she believed this would have a positive impact on student outcomes.

Student Progress

The progress students made was another factor that contributed to teachers sustaining the changes they made. Where teachers saw the differences in students' outcomes, they were more likely to continue with the changes.

Summary

Important changes occurred for teachers in both their knowledge and practice over the intervention.

Prior to the intervention, teacher knowledge of linguistic constructs was moderate, but increased impressively to a higher average of 75%. The initial lower level of knowledge has been shown as common in both overseas studies and in the two cohorts in this study.

Teachers' self-evaluation of their knowledge favoured the comprehension aspects of teaching reading prior to the intervention. Post-intervention, teachers' self-evaluation of the phonics side of teaching reading was in balance with the comprehension aspect. We believe this is significant for changes to teaching practice.

The data on teaching practice showed a change from practice that was incidental for teaching students the code, to practice that included an explicit code teaching approach. The results from both the observations rubric and the prompts tool showed a move towards a focus on the detail of the print for success in reading.

The changes in teacher knowledge and practice appear to have had an impact on the teaching for beginning readers. Increases in teacher knowledge will mean that teachers have more skills to teach students about the code explicitly and systematically. Changes to practice from implicit to explicit in approach, and a move towards directing students to a word cue first rather than meaning first approach appears to have had an impact on positive changes to student outcomes that did not occur for the comparison group. The student results show that the intervention, with a focus on teacher knowledge and explicit practice, has had a positive impact on student outcomes.

Conclusion

The results from the teacher data have implications for the teaching of reading in New Zealand.

Teachers have generally had a lack of specific training in knowledge about the English language and a corresponding lack of understanding of how to teach beginning readers in the written code of English. A lack of specific knowledge means teachers cannot be expected to teach in a systematic and explicit way. One issue may be that most initial teacher education programmes do not address the knowledge gap, possibly from a lack of time available or because of a commonly-held belief that students are best to learn the code from seeing words in a connected text (reading), or producing their own written versions of words (writing).

As indicated throughout the report, there has been a long-term emphasis on an incidental and multiple cues approach to teaching students to read. The dominant approach has remained despite extensive international and New Zealand research that has shown the importance of a focus on teaching the code to students. Along with this research, student literacy data show that many students are struggling with learning to read with the current teaching approach.

A range of resources for teacher and student use has promoted an approach that focuses on students developing a reading processing system. The system involves the

teaching of multiple cues as the way to process written text and as this report has outlined, such a process has made it difficult for some students to learn to read effectively and for teachers to have an effective way to teach.

The book series provided to teachers requires and promotes an incidental approach to teaching reading. The natural language of the texts requires that teachers promote a meaning or sentence structure approach to reading a sentence, with too many words being unknown for students to apply decoding strategies. The lesson observations showed that adherence to natural language texts for teaching early reading placed pressure on teachers and learners. Teachers are required to find on-the-spot solutions to a range of orthographic word patterns and students are required to use large amounts of working memory on the range of patterns. Early reading teaching would be enhanced by the use of decodable texts that support a developmental scope and sequence. Such texts provide support for both teachers and learners.

As a result of the intervention in the Early Literacy Project, most teachers changed their practice to place more emphasis on teaching students to use the details in the printed word. However, some teachers found it difficult to reposition the dominant focus on meaning. Teachers aligned a focus on meaning with a love of books and implied that a focus on code would mean the opposite. An idea that students would lose an excitement for reading if the focus was on the code was prevalent. Such a premise is a common argument against a focus on phonics but this is a misguided argument that lacks research support. A focus on phonics does not mean no focus on meaning or comprehension.

The Early Literacy Project placed a focus on the use of children's literature in an interactive read aloud approach for the development of vocabulary and comprehension. However, for small group reading instruction, the focus on code knowledge first and foremost has been a significant change for teachers' practice.

Some teachers commented that they first needed to know that changing practice would make a positive difference before they could be justified in making changes to practice. The same demands have not generally been placed on the current system, despite evidence that a number of students are struggling to make expected progress.

As a result of examining the teacher data and the changes in student outcomes, we believe a change in the positioning of the teaching of reading in the first year at school is

necessary. We suggest an examination of international studies and their recommendations (Hempenstall, 2016; Moats, 2009; National Reading Panel, 2000; Rose, 2006) to align teaching practice in New Zealand with the established research findings. The Implementation part of the report contains more detail of such recommendations.

OBSERVATIONS FROM THE COACH

The Role of the Independent Coach

As part of the Early Literacy Project, an independent literacy expert undertook a series of in-class literacy lesson observation visits. Intervention teachers in Cohort 2 were the focus of coaching support associated with the literacy lesson observations. The purpose of the observation visits was as follows:

- To observe New Entrant/Year 1 Intervention teachers teaching literacy-related lessons to either whole classes, small groups, or with individual students.
- To offer guidance, advice and support to the teachers related to aspects of either the lessons that were observed and/or to the literacy project in general.

Lessons were observed to gauge the extent to which teachers were able to implement teaching strategies based on information from the PLD workshops. More importantly, either during the lessons or immediately following, the Coach was able to discuss or explain relevant protocols that may have further enhanced the effectiveness of the particular lesson that was observed. The Coach reported that in some instances, full lessons were modelled to further demonstrate effective literacy practice for the teacher.

Between mid-May and mid-November 2016 the Coach completed 112 in-class literacy lesson observation visits to 28 teachers in 12 schools across the lower North Island, from New Plymouth to Wellington.

Each Intervention teacher received four separate in-class observation visits. The average duration of each visit was 45 minutes. The Coach reported that this time was frequently exceeded as further post-observation discussions took place during intervals or lunch periods. Often, the teacher was also released from their class following the lesson observation to allow further discussions.

Key Points Noted by the Coach

A number of important observations were noted by the Coach and passed on to members of the Project Team. The most notable of these are as follows:

- ***Use of phonics programmes.*** Every teacher who was observed was familiar with at least one commercial phonics-related programme. It was apparent that they had been using such programmes prior to the Early Literacy Project. Some teachers were using more than one phonics programme and most were also using supplementary phonics-type computer programmes within their daily literacy lessons.
- ***Widespread confusion in using phonics programmes.*** Often teachers were confused about how to effectively use phonics programmes. Different programmes tend to emphasise different aspects and approaches for having students learn important phonics subskills. Teachers seemed more comfortable with using a separate and structured phonics programme than with including phonics in small group instruction.
- ***Little assessment of phonic subskills.*** Although there was strong evidence that phonics-based skills and knowledge were being taught in many different forms, there was little evidence that the ongoing mastery of these subskills was being regularly monitored on an individual student basis. Apart from letter names, and to a lesser extent, letter sounds, very few teachers were able to produce evidence of having regularly monitored each student's individual progress. In general, teachers indicated that they assumed that all students learned the required subskills simply by virtue of group or whole class exposure.
- ***Limited understanding of phonological and phonemic awareness.*** Although teachers demonstrated increased understanding of the importance of phonological and phonemic awareness in early literacy acquisition, observations revealed that few teachers collected individual student data for either the initial assessment or for ongoing monitoring of the learning progress of these subskills. This occurred despite all teachers having been provided with a personal copy of relevant and easy-to-administer assessments that included basic phonological awareness knowledge. This observation highlights challenges presented for changes in teaching practice when

assessment schedules are typically well established in schools. Teachers noted that finding time to add more assessments was difficult. In line with this, some teachers explained that the New Zealand *Year 1 Literacy Curriculum* booklet was at odds with the Project PLD materials and advice because of the statement that phonemic awareness skills "do not need to be taught as they develop in conjunction with learning to read itself" (p.7). This statement is incorrect and appears to have discouraged teachers from actively implementing the strategies presented and discussed in the PLD workshops. Although more advanced phonological awareness skills can develop in conjunction with learning to read, the earlier skills still need to be taught, explicitly in most cases.

- ***Ongoing preference for context-based word identification cues.*** Observations of guided reading instruction showed teachers used a wide range of instructional strategies when students encountered problematic words. There were numerous examples of teachers encouraging students to "sound the word out", or to identify particular letter sounds or orthographic units. This finding is consistent with the data presented in the previous section on Teacher Data. However, despite the strong emphasis on the use of word-level decoding strategies in the PLD workshops, the Coach noted that teachers continued to encourage students to used context-based cues as the primary strategy for identifying unfamiliar words in text. There were many occasions where the teacher encouraged students to identify the initial sound in an unknown word, but beyond this the "sounding out" strategy frequently stalled. Some teachers used an effective decoding strategy, such as writing the word/unit on a whiteboard to highlight the relevant sound(s). However, the preferred instructional response for many teachers was to either refer the reader to the illustration, prompt about the meaning, or tell the word and move on. This observation is consistent with findings reported in this report on the teacher prompts.
- ***Post-lesson follow-up activities.*** One of the key components of an effective instructional lesson is the quality and purpose of follow-up activity. Such activities act as both a form for revision and allow students to consolidate new learning. The Coach noted that follow-up activities were seldom assigned following the literacy lessons that were observed. Where follow-up activities were assigned, they often involved tasks that did not seem to ask students to apply their new learning, or

somewhat unrelated tasks, such as colouring in letters or words, or pictures. The Coach provided examples of such activities that would assist in consolidating learning from the literacy lesson. It was noted that in many cases, time spent on colouring activities was greater than time allocated to word-learning tasks.

- ***Assessment practices.*** Despite every teacher receiving an Assessment Handbook and having experience in the use of assessments as part of the PLD workshops, the use of key assessments related to foundational knowledge was markedly variable. As noted earlier in this section, there was little evidence from many teachers that they were systematically assessing students' acquisition of foundational skills and knowledge. Most of the essential knowledge and subskills for successful early literacy development are finite in number and quantifiable (e.g., letter names and sounds, basic phonological and phonemic awareness, blends and digraphs). This means that rate of learning progress for each student can be accurately assessed in an objective and precise manner, using assessments provided to each teacher by the Project.
- ***Role of coaching in teacher change activities.*** The Coach reported that in general, teachers were very appreciative of the demonstrations and discussions. We are confident that inclusion of a coach/mentor/observer added a positive element to the Early Literacy Project. There is value in including such a role in future professional development initiatives that involve working with teachers, especially where there is a focus on changing instructional behaviours and beliefs.

CHALLENGES

During the course of the project, we encountered a number of unanticipated challenges. These are summarised as follows:

- Obtaining a reasonably large number of schools to enable use of a randomised control design proved to be unattainable.
- Despite explanations about the need for Intervention students to remain in the same class with their teacher who participated in the PLD workshops, a number of schools were unable to accommodate this request.
- Although we anticipated that Intervention teachers would take time to "digest" the information and practical activities presented in the PLD workshops, it appears from the examination of Cohort 1 teaching video data that more explicit guidance in these workshops was required.
- Our attempts to establish an online forum for teachers to exchange ideas and experiences were unsuccessful. Our model for this approach was the largely successful online forum (based on the Moodle learning management system) that forms an integral part of post-graduate Education coursework at Massey University.
- Requesting teachers to complete an online survey and assessment of their knowledge of the language foundations associated with literacy instruction met with a very mixed response. Despite requests to answer all questions (even with a "Don't Know" response) and to complete all sections of the survey, many questions were left unanswered and some sections completely overlooked. Based on these experiences with Cohort 1 teachers, we changed the approach to obtaining participation in the teacher survey and linguistics knowledge questionnaire from online delivery to pencil-and-paper delivery within PLD workshop sessions. This change led to an improved response rate.
- It proved to be very challenging in some cases to obtain information from some schools regarding student backgrounds (ethnicity, parental occupations, student gender in some cases), and reading book levels. Despite repeated requests, some schools did not provide important data for students.

- The resources available to teachers present a challenge, and indeed are a barrier, to changing practice. The predominant book series used in classrooms follows a natural language (*Ready to Read*) or controlled reading vocabulary (Price Milburn, *PM+*) approach. The types of word patterns that occur in such texts are diverse. Natural language texts constrain teachers to a particular way of teaching. The Teacher Support Materials that accompany the *Ready to Read* series give guidance on a multiple cues approach with very little if any guidance for teachers on the code knowledge needed in reading. The materials run counter to attempts in this project to develop in teachers a range of strategies for enhancing word-level decoding skills in their students. Such texts also result in very limited benefits from the widespread use of commercial phonics programmes.

LIMITATIONS OF THE RESEARCH

Research Design

We planned in this study to adopt a mixed methods randomised control design with a longitudinal component. We intended to include in the study approximately 1,600 new entrants who were starting school for the first time in February 2015. To obtain this total number of students we randomly selected 80 schools in the lower North Island that in total would likely enrol this number of New Entrant students at the start of 2015.

By the start of the 2015 school year, less than half the 80 schools approached (39) confirmed their willingness to take part in the research.

The number of New Entrant students from these schools totalled 359, which was well short of our planned 800 students each in the Intervention and Comparison groups.

As a result, instead of having a randomised control design for the research, we were left with a quasi-random volunteer sampling design. In some of our analyses, cell sizes in the Group by Decile Band design were smaller than desired. Very small numbers of students from most ethnic backgrounds also meant that analyses of variables that included ethnicity were not appropriate.

Incomplete Information from Schools

As we have indicated in various parts of this report, we were unable to obtain all information from schools regarding ethnic background, parental occupations, and perhaps more importantly, reading book levels. Incomplete information reduced the sample sizes for some of the analyses.

Teacher Changes During Year 1

At the start of the project we emphasised to schools the importance of students remaining with teachers who participated in the PLD workshops. The purpose of this request was to assess the impact on students' literacy learning outcomes as a function of PLD participation, compared to students in the Comparison schools whose teachers did not participate in the workshops. This request was unable to be met by a number of schools, as reported in regard to the Cohort 1 study.

STRENGTHS OF THE PROJECT

Despite the challenges and limitations in the research, there are a number of strengths that should be noted.

Theoretical Basis

A major strength of this project is the strong, contemporary theoretical basis that guided the approach to literacy acquisition and literacy instruction. We adopted the Cognitive Foundations of Learning to Read framework (Tunmer & Hoover, 2014; see), which combines the cognitive elements underpinning the development of the language comprehension and word recognition components of the Simple View of Reading (Gough & Tunmer, 1986). It is based on extensive research spanning nearly three decades showing that learning to read follows a developmental progression from pre-reader to skilled reader involving qualitatively different but overlapping phases.

The theoretical basis of this project emphasised the importance of the development of word-level decoding skills as a necessary, though not sufficient, requirement for developing reading comprehension skills. In other words, reading for meaning, the primary goal of reading, involves children being able to quickly decode words, especially those words that carry the meaning in connected text.

Research Design

Although we were unable to implement a truly randomised control research design, nonetheless there are a number of design strengths.

The overall sample size, involving 729 students from 38 schools, is representative of a wide range of decile rankings and school sizes. Results from both cohorts, especially the significantly positive results from Cohort 2, are reasonably generalizable.

Although not part of the initial plan for the project, having two cohorts, with two Intervention and two Comparison groups, has provided a good basis for making comparisons about outcomes associated with the teacher PLD workshops. In Cohort 1, the results were disappointing. Following changes to PLD delivery, the Cohort 2 Intervention group achieved positive results that are consistent with theory and previous research.

The Role of a Coach

Having a coach support teachers, and in some cases, providing "demonstration" lessons, added an important element to the project. Moreover, the adoption of a coach reinforced the importance of ongoing support for teachers engaged in PLD that is designed to change aspects of instructional practice. Most teachers expressed gratitude for the support of the coach and adopted many practical suggestions made following observation sessions.

SUMMARY AND CONCLUSIONS

Summary

Two cohorts of teachers and students took part in this research project. Cohort 1 students commenced school in February 2015. Literacy-related assessments were collected over 2 ½ years from school entry to the middle of 2017, when the students were in Year 3. Cohort 1 teachers took part in the PLD workshops during 2015. Cohort 2 students started school in February 2016. Literacy assessments were collected from school entry to the middle of 2017, when the students were in Year 2. Cohort 2 teachers participated in PLD workshops during 2016. Each cohort had a number of comparison schools, with teachers who continued with their normal literacy instruction. Students in these schools were also assessed.

Results for Cohort 1 students showed that the intervention did not have a positive impact on student literacy learning outcomes. Intervention students did not outperform Comparison students on any of the literacy assessments over the 2½ years of the project. Comparison students generally performed better than the Intervention students. Further, teachers showed no evidence of change in their knowledge of the language foundations associated with literacy teaching and learning, although changes in instructional practices were clearly made by some teachers. We attributed these findings to unexpected challenges faced by teachers in implementing new teaching strategies presented in the PLD workshops.

Results in the latter part of 2015 showing that Intervention students were not improving relative to Comparison students led to modifications to the PLD delivery and supporting materials. A second cohort of teachers and students was recruited from the pool of 38 schools that remained in the project.

Cohort 2 Intervention students showed significantly better literacy learning outcomes than the Cohort 2 Comparison students at the end of Year 1 (2016) and in the middle of Year 2 (2017). Cohort 2 Intervention students also significantly outperformed Cohort 1 Intervention students at the mid-Year 2 assessment point on measures of reading and spelling. Especially significant was the finding that low decile Intervention students dramatically outperformed low decile Comparison students, and in some cases had mean

literacy assessment scores that were close to or equal to those of students in higher decile schools.

Cohort 2 Intervention teachers showed important improvements in their knowledge of the language foundations associated with effective literacy teaching and learning. Video samples of classroom teaching also revealed changes in instructional practices that reflected content and materials from the PLD workshops.

In terms of the four specific research questions, none of the research questions was answered in the affirmative for Cohort 1 students and teachers. For Cohort 2 however, three of the four research questions were answered in the affirmative:

1. Cohort 2 Intervention students improved their literacy learning outcomes compared with students in the Comparison group. Results for reading and spelling in the middle of Year 2 (2017) were particularly impressive indicators of the benefits of the intervention.
2. The proxy indicator of motivation, reading self-efficacy, did not reveal differences between the Intervention and Comparison groups. This is likely due to the finding that this assessment failed to adequately discriminate between higher and lower literacy achievers. Particularly unexpected was the finding that the low decile Comparison students had high (positive) self-efficacy scores, despite their literacy achievement scores on most measures being very low. Discrepancies like this are possibly due to unrealistic self-perceptions forming as a result of non-task specific, positive teacher feedback that does not relate to actual task performances.
3. For Cohort 2 low decile Intervention students, the intervention provided compelling evidence that the literacy achievement gap was significantly reduced, and in some cases almost closed.
4. Cohort 2 Intervention teachers revealed through improved knowledge of basic language constructs, improved self-perceptions of literacy teaching, and through observations of teaching practice that they had generally increased in both their ability and confidence in teaching word-level skills.

Conclusions

This project was undertaken to trial literacy teaching strategies that would contribute to an improvement in the literacy performance levels of New Zealand students, especially during the foundation New Entrant/Year 1 period of schooling.

Our approach to the teacher PLD activities was based on the scientific research on how children learn to read, which indicates that the ability to read for meaning (and enjoyment), depends on the ability to quickly and accurately recognise words in text. When high levels of automaticity in word recognition occur, cognitive resources can be allocated to sentence comprehension and text integration processes (Pressley, 2006), that is, to the *meaning* of text. Students who experience cognitive overload because of difficulty in word identification struggle to obtain the meaning of sentences in text, because it is the unfamiliar words that typically convey the meaning.

We discussed with Intervention teachers that the essential requirement for discovering the mappings between spelling patterns and sound patterns is the ability to segment spoken words into subcomponents. Young students who have ongoing difficulties in detecting phonemic sequences in words (i.e., have low phonemic awareness) are unable to fully grasp the alphabetic principle and discover spelling-to-sound relationships (e.g., Shankweiler & Fowler, 2004). Understanding the alphabetic principle, or "cracking" the alphabet code, is necessary (but not sufficient by itself) for being able to read for meaning.

The teacher PLD programme focussed on developing in teachers the high level of teacher knowledge that is required for effective literacy teaching. Teacher knowledge of English orthography and morphology makes it easier for teachers to understand patterns of word decoding and word spelling. In turn, this understanding helps them to assist students in learning the essential skills for reading and spelling (McNeill & Kirk, 2013).

The Cognitive Foundations for Learning to Read framework (Tunmer & Hoover, 2014; Figure 1 in this report) provided the conceptual basis for the project. We added the key component to PLD workshop materials that revolved around the use of explicit, structured and systematic instruction. For at least 30 years, this approach has generally been eschewed in New Zealand literacy teaching in favour of a more implicit and incidental approach to learning to read. Drawing from extensive research, we sought to develop in Intervention teachers the understanding that effective early literacy instruction should include the

elements of each letter name and sound, blending of phonemes together to decode unfamiliar words in their entirety, segmenting sounds in spoken words to spell unknown words, as well as teaching the meanings of target words. In short, knowledge of the specific phonic elements required for effective word learning is necessary for progress to be made with reading acquisition.

Although our approach was unsuccessful with teachers and students in Cohort 1, results for Cohort 2 teachers and students provide evidence that a more systematic and explicit approach to the teaching of word-level decoding skills is beneficial. Our findings in this regard are consistent with theory and previous research (e.g., Hattie, 2009; Snow & Juel, 2005; Tunmer & Arrow, 2013). The results are especially beneficial for students attending low decile schools.

We made a number of observations during the project relating to challenges teachers faced in adopting and implementing teaching strategies that were often quite different from their typical approach to literacy instruction. Teachers in both cohorts were generally positive about the opportunities to participate in the project and to attend the PLD workshops. Feedback and observation of teachers teaching, however, revealed the difficulties in changing deeply embedded practices.

Observations, interviews and feedback revealed that a major barrier to teachers changing their practice was the strong belief that learning to read involves a multiple cues approach for the identification of unfamiliar words in text, rather than a process which of necessity involves children “cracking the code” in order to derive meaning from text. This multiple cues approach has formed the basis of reading instruction in initial teacher education programmes in New Zealand for over 30 years. Materials that teachers use for teaching practice, guidance, and implementation promote the multiple cues approach (e.g., *Effective Literacy Practice in Years 1-4,* teacher support materials with the *Ready to Read* texts). The multiple cues approach means teachers emphasise that children use meaning and sentence structure to enable the reading process. Word level detail is used only as necessary and often as a last resort (Clay, 2005; Goodman, 2008). Switching to a systematic and explicit approach, with a greater emphasis on the development of word-level decoding skills, proved difficult for many teachers, even though most saw the merits of such a change.

More intensive PLD over a longer period of time would have been beneficial for most teachers, with probably stronger outcomes for students.

Similarly, set teaching practices emerged as a constraint for change in practice. On pre-intervention surveys, teachers favoured a guided reading approach where code knowledge was taught incidentally during the reading of a levelled book. Changing the nature of guided reading to include more explicit and systematic instruction on code knowledge was difficult for some teachers. Many video clips showed teachers successfully adding an explicit teaching section on code knowledge to their lessons and attempting to follow the scope and sequence to teach children about the code. However, some teachers struggled to make this change and remained reliant on a guided reading approach and the selected levelled book for incidental instruction on the code.

Data from the project indicate that part of the difficulty many teachers had in adopting a more comprehensive, explicit and systematic approach to teaching code knowledge is due to the resources that most schools use to support early literacy instruction. The *Ready to Read* book series teachers commonly use in New Entrant/Year 1 classrooms requires and promotes an implicit or incidental word identification learning strategy. The natural language of the texts is designed to focus on meaning or sentence structure. As a result, too many complex and unknown words obstruct students' efforts to apply decoding strategies. Video clips of literacy lessons showed teachers and learners frequently struggling with natural language texts because of the wide range of orthographic word patterns. Such a range requires students to use a considerable amount of working memory to deal with the range of patterns. Rather than making the process of reading easier for students, the videos showed evidence that they were highly dependent on the teacher in reading the text. Early reading teaching would be enhanced by the use of decodable texts that support a developmental scope and sequence. Such texts provide support for both teachers and learners.

Although we have drawn attention here to some important barriers to changing instructional practices, there were many enablers that we have discussed earlier in the report. Clearly, the enablers (e.g., improved teacher knowledge and student learning; scope & sequence guidance; provision from the project of decodable readers) outweighed the barriers to change, as shown in the significant literacy learning outcomes for Cohort 2 students.

As a result of examining the teacher data and the changes in student outcomes, a change in the positioning of the teaching of reading in the first year at school is necessary. We suggest an examination of international studies and their recommendations (e.g., Hempenstall, 2016; Moats, 2009; National Reading Panel, 2000; Rose, 2006) to align teaching practice in New Zealand with the established research findings.

IMPLICATIONS AND RECOMMENDATIONS

Implications

The results of this research project indicate that teachers can adopt more effective literacy instructional strategies for the benefit of all students, and especially those students in low decile schools. This is a positive outcome. Changes made to literacy instruction in New Entrant/Year 1 classes have the potential to reduce the literacy learning gap between good and poor readers, and in turn improve New Zealand's overall literacy performance.

It is well known that students in the long tail of poor literacy performance tend to come from disadvantaged background, and tend to include disproportionately larger numbers of Māori and Pasifika students (Nicholson, 1997; Tunmer & Chapman, 2015; Tunmer, Chapman, & Prochnow, 2003, 2006). Children from disadvantaged backgrounds are much more likely to have low literacy on entry to school (Prochnow, Tunmer, & Arrow, 2015). Unless effective Year 1 teaching ameliorates these school entry differences, difficulties in learning to read result in difficulties with reading to learn, thereby perpetuating a cycle of low educational attainment and (often) poverty.

We are not dismissing the non-educational factors that cause and perpetuate disadvantage and poverty in New Zealand. However, there is one domain in which education can play a significant role in providing children from disadvantaged backgrounds with genuine equality of educational opportunity. Improved literacy instruction can provide a positive impact on learning opportunities during the early years of schooling (Hempenstall, 2016).

Major reviews of research on reading over at least the last decade agree on the key elements of effective reading programmes and the most effective way of teaching them (Hempenstall, 2016). The most comprehensive review was published by the National Reading Panel (NRP) in the United States (National Reading Panel, 2000). Five key areas were identified by the NRP as essential for effective, evidence-based reading instruction:

- Phonemic awareness
- Phonics
- Vocabulary
- Fluency
- Comprehension

The first three elements are especially important during the first year of formal reading instruction.

The review and recommendations of the NRP are consistent with the findings of other reviews and international reports, including the Rose Report (2006) in the UK, the UK Primary National Strategy (2006), and the National Inquiry into the Teaching of Literacy (2005) in Australia. All reports emphasised the importance of explicit and systematic teaching of the components associated with the alphabetic code. The Australian Inquiry into the Teaching of Literacy (2005) summarised this key point as follows:

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

Before proceeding further, it's important at this point to clarify the confusion that typically surrounds the term *phonics*. Hempenstall (2016) notes that phonics has several related meanings:

- The relationship between the sounds in speech and the symbols in written language that represent those sounds;
- The methods used to teach those relationships;
- The phonological process of using the relationship to decode (sound out) a new word.

Teaching phonics should occur early. Phonics is essential for struggling readers, students with reading disabilities, students with intellectual disabilities, and English language learners (Hempenstall, 2016). The vast majority of schools in New Zealand use commercial phonics programmes with students in New Entrant/Year 1 classes (Chapman et al., 2017). There is evidence, however, that teachers have not received adequate training in the effective use of

phonics as an *integral* part of literacy instruction. Further, the wide range of phonics programmes in use in New Zealand (over 50 commercial programmes; Chapman et al., 2017) indicates significant misunderstandings about what constitutes an effective, research-based phonics programme.

More importantly, phonics instruction is typically based on *analytic phonics,* in which phonics concepts are taught as they arise in natural language texts. Another common New Zealand approach is to use commercial phonics programmes in a stand-alone manner, outside of the act of reading itself. Finally, phonics is often taught as spelling rather than as a reading strategy as well. The different findings for Cohort 1 and 2 teachers and students reflect the change of emphasis in the PLD programme to explicitly supporting teachers to adopt systematic, explicit instruction that is supported by decodable texts, rather than phonics driven by natural language texts.

The results of our research for Cohort 2 students are consistent with the extensive research on effective literacy teaching and learning. To embrace the findings from this project and from the extensive reviews of literacy research elsewhere, changes to literacy instruction and supporting materials will be needed in New Zealand. Accordingly, we make the following recommendations.

Recommendations

1. Teacher knowledge of the nature of language is essential for teachers to provide effective instruction in the foundations of reading. Just as other professions require technical understanding of foundational knowledge (e.g., doctors, lawyers, engineers, accountants), so too do teachers. As Moats (2009) asserted, teachers need to be analytical about language and need such knowledge to systematically and explicitly teach foundational skills to beginning readers. Results from this project demonstrate the positive effects on students' literacy learning outcomes when teachers increase their knowledge of language foundations.

    A national strategy should be developed to upskill New Entrant/Year 1 teachers in the importance, knowledge, and use of

foundational language skills involved in successful literacy learning.

2. The fundamental nature of literacy resources for teachers, and the philosophy that underpins them, have remained largely unchanged over the past 30 years. Attempts to use commercial phonics programmes in conjunction with these resources is largely fruitless. As shown in this project, attempts to use systematic and explicit instruction for developing effective word decoding skills are compromised by the use of levelled readers that are based on "natural language". Similarly, instructional resources used by most New Zealand classroom teachers of New Entrant/Year 1 students are significantly inconsistent with the last two decades of research on literacy instruction.

   The instruction guidebook "Effective Literacy Practice in Years 1 to 4" should be phased out and replaced by a much more contemporary text for teachers, based on the abundance of contemporary research frequently mentioned in this report.

   Teachers cannot be expected to make necessary changes to literacy instruction without an appropriately up-to-date, research-led text to support new practices.

3. We found during the course of our project that all participating teachers wanted to be more effective in their literacy instruction. Most saw value in and understood the material presented in PLD workshops about the importance of beginning readers developing effective word decoding skills. And most valued the opportunity to learn about the importance of understanding the language foundations that underpin effective literacy teaching and learning. Implementing learnings from the PLD workshops and accompanying materials was more challenging. The engagement of a coach helped. To bring about significant change in literacy teaching practices will require a systematic and sustained approach to in-service PLD for upskilling teachers' knowledge and instructional practice.

A strategy should be developed for the implementation of a comprehensive PLD programme designed to provide teachers of New Entrant/Year 1 students with effective tools for teaching the five key areas required for effective literacy instruction.

4. During the course of the project, teachers often remarked that material covered in the PLD sessions was clearly important, but never referred to in initial teacher education literacy programmes. Similar remarks from post-graduate students in literacy courses at Massey University frequently drew attention to the shortcomings of initial teacher education programmes. There are many reasons for this. However, New Zealand colleges/faculties/schools of education should ensure that literacy courses adopt approaches that are based on contemporary scientific research.

A process for instituting change in initial teacher education literacy courses should be developed and implemented.

It is a truism and an indicator of insanity, that if we continue to engage in the same teaching practices we will ensure the same (or worse) learning outcomes. The findings from this research project, despite some limitations, provide evidence in support of a more effective approach to literacy instruction in New Zealand New Entrant/Year 1 classrooms. Students, parents, and the New Zealand community at large, deserve to have beginning readers receive the very best, research-led literacy instruction possible.

REFERENCES

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

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 and education* (pp. 241-249). London: Routledge.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.). *Self-efficacy beliefs of adolescents*, (Vol. 5., pp. 307-337). Greenwich, CT: Information Age Publishing.

Binks-Cantrell, E., Washburn, E. K., & Joshi, R. M. (2012). Validation of an instrument for assessing teacher knowledge of basic language constructs of literacy. *Annals of Dyslexia, 62,* 153-171.

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: Psychology Press.

Braid, C. (2012). *Children's responses to a picturebook during a small group, co-constructed read-aloud.* Master's thesis, Massey University, Palmerston North, New Zealand.

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

Chapman, J.W. (2016). *Research report on the predictive abilities and effectiveness of assessment tools for monitoring literacy programmes in primary schools, involving the Quick60 programme.* Palmerston North: Massey University. DOI: 10.13140/RG.2.1.1962.9046

Chapman, J.W., Allcock, J., & Cody, J. (2016). *Language to literacy: The Shine literacy project*. Porirua, New Zealand: Shine Literacy Foundation.

Chapman, J.W., Arrow, A.W., Tunmer, W.E., & Greaney, K.T. (2015). *Survey of teacher literacy knowledge and efficacy: A report of findings for the Ministry of Education.* Palmerston North: Massey University.

Chapman, J.W., Greaney, K.T., Arrow, A.W., & Tunmer, W.E. (2017, July). Phonics Use, Teachers' Knowledge of Language Constructs, and Teachers' Literacy Teaching Practices in New Zealand: Prospects for Struggling Readers in a Predominantly Whole

Language Instructional System. Paper presented at the annual meeting of the International Academy for Research in Learning Disabilities, University of Queensland, Brisbane, Australia.

Chapman, J.W., Greaney, K.T., & Prochnow, J.E. (2015). Literacy performances of young adults: Outcomes of school-based literacy instruction. In W.E. Tunmer & J.W. Chapman (Eds.), *Excellence and equity in literacy instruction: The case of New Zealand* (pp. 71-92). Basingstoke, England: Palgrave Macmillan.

Chen, D., Hu, B. Y., Fan, X., Li, K. (2014). Measurement quality of the Chinese Early Childhood program rating scale: An investigation using multivariate generalizability theory. *Journal of Psychoeducational Assessment, 32,* 230-248.

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

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

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

Connor, C. M. (2013). Commentary on two classroom observation systems: Moving toward a shared understanding of effective teaching. *School Psychology Quarterly, 28,* 342-346.

Connor, C. M., Spencer, M., Day, S. L., Giuliani, S., Ingebrand, S. W., McLean, L., et al. (2014). Capturing the complexity: Content, type, and amount of instruction and quality of the classroom learning environment synergistically predict third graders' vocabulary and reading comprehension outcomes. *Journal of Educational Psychology, 106,* 762-778.

Crawford, A. D., Zucker, T. A., Williams, J.M., Bhavsar, V., Landry, S. H. (2013). Initial validation of the pre-kindergarten classroom observation tool and goal setting system for data-based coaching. *School Psychology Quarterly, 28*, 277 -300.

Crooks, T., & Caygill, R. (1999, December). New Zealand's National Monitoring Project: Maori 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.

Crystal, D. (2003). *The Cambridge encyclopedia of the English language*. Cambridge, England: Cambridge University Press.

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.

Davis, A. (2007). *Teaching reading comprehension*. Wellington, New Zealand: Learning Media.

de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6*, 51-77.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*, 81-112. doi: 10.3102/01623737024002081

Doabler, C., & Nelson-Walker, N. J. (2009). *Ratings of classroom management and instructional support*. Eugene, OR: Center on Teaching and Learning, University of Oregon.

Duffy, G. G. (2009). *Explaining reading*. New York: Guilford Press.

Dunn, L. M., Dunn, D., M, Sewell, J., Styles, B., Brzyska, B., Shamsan, Y., et al. (2009). *British Picture Vocabulary Scale* (3rd ed.). London: GL Assessment.

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.

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 18*, 1, 5-21.

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

Fielding-Barnsley, R. (2010). Australian pre-service teachers' knowledge of phonemic awareness and phonics in the process of learning to read. *Australian Journal of Learning Difficulties, 15*, 99-110. doi: 10.1080/19404150903524606

Finn, P.J. (1977). Word frequency, information theory, and cloze performance: A transfer feature theory of processing in reading. *Reading Research Quarterly, 13,* 508-537.

Flockton, L., & Crooks, T. (1997). *Reading and speaking assessment results 1996: National Education Monitoring Report 6.* Wellington, New Zealand: Ministry of Education.

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading : Good first teaching for all children*. Portsmouth, NH: Heinemann.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*, 915-945. doi: 10.3102/00028312038004915

Gersten, R., Compton, D. L., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to intervention and mulit-tier intervention in the primary grades* (No. NCEE 2009-4045). Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences, U.S Department of Education. Retrieved from http://ies.ed.gov/ncee

Gilmore, A.M. (1998). *School entry assessment*. Wellington, New Zealand: Ministry of Education.

Gluckman, P. (2013). *The role of evidence in policy formation and implementation: A report from the Prime Minister's Chief Science Advisor.* Auckland, New Zealand: Office of the Prime Minister's Science Advisory Committee.

Goodman, K. E. (1976). Behind the eye: What happens in reading. In H. Singer and R. B. Ruddell (Eds.), *Theoretical models and processes of reading*, 2nd ed. (pp. 470-496). Newark: DE: International Reading Association.

Gough, P.B. (1983). The beginning of decoding. *Reading and Writing, 5,* 181-192.

Gough, P.B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6-10.

Greaney, K. (2001). An investigation of teacher preferences for word identification strategies. *Australian Journal of Language and Literacy, 24*, 21–30.

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

Hempenstall, K. (2016). *Read about it: Scientific evidence for effective teaching of reading*. Sydney, Australia: Centre for Independent Studies. https://www.cis.org.au/app/uploads/woocommerce_uploads/2016/03/rr11.pdf

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

Mather, N., Bos, C., & Babur, N. (2001). Perceptions and knowledge of preservice and inservice teachers about early literacy instruction. *Journal of Learning Disabilities, 34,* 472-482.

McKay, M. F., & Barnard, J. (1999). *Neale Analysis of Reading Ability* (3rd ed.). Melbourne, Australia: Australian Council for Educational Research.

McNeill, B., & Kirk, C. (2014). Theoretical beliefs and instructional practices used for teaching spelling in elementary classrooms. *Reading and Writing*, 27, 535-554.

Ministry of Education. (1997). *Reading and beyond*. Wellington, New Zealand: Learning Media.

Ministry of Education. (2003). *Effective literacy practice in years 1 to 4.* Wellington, New Zealand: Learning Media.

Ministry of Education. (2011). *Ministry of Education briefing to the incoming Minister*. Wellington, New Zealand: Ministry of Education.

Moats, L. (1994). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia, 44*, 81–102.

Moats, L. (1999). *Teaching reading IS rocket science: What expert teachers of reading should know and be able to do*. Washington, DC: American Federation of Teachers.

Moats, L. (2009). Still wanted: Teachers with knowledge of language. *Journal of Learning Disabilities 42*, 387-391.

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

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

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*

reports of the subgroups (No. BBB35631). Bethesda, MD: National Institute of Child Health and Human Development. www.nichd.nih. gov/publications/pubs/nrp/documents/report.pdf

Nicholson, T. (1991). Do children read words better in context or in lists? A classic study revisited. *Journal of Educational Psychology, 83*, 444-450.

Nicholson, T. (1993). Reading without context. In G. B. Thompson, W. E. Tunmer & T. Nicholson (Eds.), *Reading acquisition processes* (pp. 105-122). Clevedon, England: Multilingual Matters.

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., & 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.

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story? *Reading and Writing, 23*, 189-208.

Pianta, R. C. , La Paro, K., Hamre, B. K. (2008). *Classroom assessment scoring system: CLASS.* Baltimore, MD: Paul H. Brookes.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 23*, 431-451. doi: http://dx.doi.org/10.1016/j.ecresq.2008.02.001

Piasta, S. B., Connor, C. M., 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.

Pressley, M. (2006). *Reading instruction that works: The case for balanced literacy instruction* (3rd ed.). New York: Guilford Press.

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

Reddy, L.A., Dudek, C.M, Fabiano, G.A., Peter,S. (2015). Measuring Teacher self-report on classroom practices: Construct validity and reliability of the classroom strategies scale: Teacher form. *School Psychology Quarterly, 30,* 513-533.

Reddy, L. A. & Dudek, C.M. (2014). Teacher progress monitoring of instructional and behavioral management practices: An evidence-based approach to improving classroom practices. *Journal of School and Educational Psychology, 2,* 71-84.

Reddy, L. A., Dudek, C. M., Fabiano, G. A., & Peters, S. (2015). Measuring teacher self-report on classroom practices: Construct validity and reliability of the classroom strategies scale: Teacher form. *School Psychology Quarterly, 30,* 513-533.

Rose, J. (2006). *Independent review of the teaching of early reading.* Annesley, England: Department for Education and Skills.

Schatz, E.K., & Baldwin, R.S. (1986). Context clues are unreliable predictors of word meanings. *Reading Research Quarterly, 21*, 439– 453.

Senechal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research*, Vol. 2 (pp. 173-182). New York: Guilford Press.

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.

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

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/pubs/monograph_reports/MR1465

Snow, C. E., & Juel, C. (2005). Teaching children to read: What do we know about how to do it. M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. Malden, MA: Blackwell Publishing.

Spear-Swerling, L., & Zibulsky, J. (2014). Making time for literacy: Teacher knowledge and time allocation in instructional planning. *Reading and Writing, 27*, 1353-1378. doi: 10.1007/s11145-013-9491-y

Spear-Swerling, L., Brucker, P. O., & Alfano, M. P. (2005). Teachers' literacy-related knowledge and self-perceptions in relation to preparation and experience. *Annals of Dyslexia, 55*, 266-296.

Stahl, K. A. D., Keane, A. E., & Simic, O. (2013). Translating policy to practice: Initiating RTI in urban schools. *Urban Education, 48*, 350-379. doi: 10.1177/0042085912451755

Stahl, S. A. (1997). Instructional models in reading: An introduction. In S. A. Stahl and D. A. Hayes (Eds.), *Instructional models in reading*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stuart, M, Masterson, J., & Dixon, M. (2000). Spongelike acquisition of sight vocabulary in beginning readers? *Journal of Research in Reading, 23*, 12-27.

Tong, X., Deacon, S.H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: another piece of the puzzle. *Journal of Learning Disabilities, 47,* 22-33. doi: 10.1177/0022219413509971.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805.

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. (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. doi: 10.1177/0022219411432685

Tunmer, W.E., & Chapman, J.W. (2015). *Excellence and equity in literacy instruction: The case of New Zealand*. Basingstoke, England: Palgrave Macmillan.

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. (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. (2010). Defining dyslexia. *Journal of Learning Disabilities, 43*, 229-243.

Tunmer, W. E., & Hoover, W. A. (2014). *Cognitive Foundations Model of Learning to Read*. Unpublished paper. Palmerston North, New Zealand: Massey University.

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.

Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography.* New York: 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., Torgesen, J.K., Rashotte, C.A., & Pearson, N.A. (2013). *Comprehensive Test of Phonological Processes,* 2nd Ed. Austin, TX: Pro-Ed.

Walpole, S. & McKenna, M.C. (2013). *The literacy coach's handbook: a guide to research-based practice.* 2nd ed. New York: Guilford Press.

Washburn, E.K., Joshi, R.M., & Binks-Cantrell, E.S. (2011). Teacher knowledge of basic language concepts and dyslexia. *Dyslexia: An International Journal of research and Practice, 17,* 165-183.

Wilkinson, G.S., & Robertson, G.J. (2008). *Wide Range Achievement Test,* 4th ed. Los Angeles: Western Psychological Services.

APPENDICES

APPENDIX 1

Time 1 Student Assessments (February/March 2015)

Letter Identification. Letter name and letter sound knowledge were assessed in terms of both upper case and lower case letters, using the Letter Identification task in the Diagnostic Survey (Clay, 1985). Students were asked to name each letter and to say the sound the letter represented for 26 upper case and 28 lowercase letters, two of which appeared in varying fonts. Scoring was based on the number of letters correctly identified by name, and by sound.

Vocabulary Knowledge. We used the British Picture Vocabulary Scale (BPVS: Dunn et al., 2009) to assess vocabulary knowledge. This knowledge refers to understanding the meaning of words, which is necessary for the production of functional language. Raw scores are converted to standard scores, which are related to the age of each participant.

Word Recognition. Word recognition refers to the fluent, rapid reading of words as they appear. Such words are usually known as *sight words*. We used one of the *Ready to Read* test lists (Clay, 2002). These tests comprise 45 words of the most frequently occurring words in the 12 "little" books of the *Ready to Read* series. We administered the first 15 words in one of the lists. Scoring was based on the number of words read correctly by each child. In addition, attempts at word reading accuracy were assessed by scoring the number of correct letter-to-sound correspondences in each word.

Invented Spelling. Invented spelling was assessed by having students write 18 words that were read aloud by the research assistant. The 26 (lower case) letters of the alphabet were displayed across the top of the students' response sheets. Each word that students wrote down received a score from 0 to 4. Maximum points were awarded if the sounds in the word were represented with letters, although unconventionally (e.g., *kik* for *kick*, *fil* for *fill,*

sid for *side*). Two points were awarded if more than one phoneme (but not all) was represented with phonetically related or conventional letters (e.g., *sd* for *side, lup* for *lump*). One point was awarded where the initial phoneme was represented with the correct letter (e.g., *f* for *fat*). Students were also asked to identify the sounds in the words that were read aloud. The total number of possible points for letters and sounds was 72 each.

Phonological Processing. Phonological processing was assessed using the Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2: Wagner, Torgesen, Rashotte, & Pearson, 2013). This test is normed in the United States for use with people from 4 years to 25 years. The CTOPP-2 is used to help evaluate phonological processing abilities as a prerequisite to reading fluency. We administered three of the subtests: elision, blending and matching. Elision measures the ability to remove phonological segments from spoken words to form other words. There are 34 items in this test, with discontinuation occurring when each child missed three consecutive items. Blending Words measures the ability to synthesize sounds to form words. There were 33 items in this section; again, discontinuation occurred following three consecutive missed items. Sound Matching measures the ability to select words with the same initial and final sounds. This section comprised 26 items; testing was discontinued following three missed items.

APPENDIX 2

Teacher Survey Time 1

Teacher Knowledge. The teacher knowledge survey was based on a measure of teachers' knowledge of basic language constructs validated by Binks-Cantrell, Joshi and Washburn (2012). Basic language constructs considered essential for early reading success include phonological and phonemic awareness, the alphabetic principle (phonics), and morphology (Binks-Cantrell et al., 2012). Based on extensive research during the late 1980s and 1990s (e.g., Adams, 1990; Moats, 1999), the National Reading Panel (2000) in the US, stressed the importance of teachers having an explicit knowledge of such concepts for the effective teaching of decoding skills in a direct, systematic way to enable the successful acquisition of early reading skills for all beginning readers (Binks-Cantrell et al., 2012).

The Binks-Cantrell et al. (2012) scale included 46 questions that examined teachers' understanding of basic language constructs in terms of knowledge and skills in relation to phonological and decoding elements. For example, the question "A phoneme refers to…" is defined as a knowledge question in relation to phonemic understanding within the phonological domain. Skill-based items, for example, required teachers to count the number of phonemes in a word, such as *moon*, as well as count the number of syllables and morphemes in words such as *observer* and *frogs*. In addition, we included items designed to assess teacher's perceived teaching ability, such as "evaluate your knowledge of teaching phonemic awareness and comprehension". Binks-Cantrell et al. (2012) reported that the teacher knowledge measure has a Cronbach's alpha coefficient of 0.90 and good construct validity.

In our survey of teacher knowledge, 38 items were categorised into phonemic, phonic, phonological, and morphological skills/knowledge. An additional 8 items involved teacher self-evaluations of their perceived literacy-related teaching ability.

Word Identification Prompt Scenarios. The word identification prompt task was based on six scenarios used by Greaney (2001). These scenarios were selected from two series of

publications commonly used in New Zealand primary schools; the *Ready to Read* series, which is used in most junior classes, and the school journals. The scenarios were selected to exemplify three main types of reading errors (Greaney, 2001). Type A reading errors include a non-verbal response from a reader when she/he comes across an unfamiliar word, or a minimal response such as the initial letter only. Three scenarios involved Type A errors. One scenario involved a Type B error, in which the reader gave a non-word response for the target word (e.g., "brost" for *breakfast*). Type C errors involved the reader providing a real-word substitution that makes grammatical sense, but which is nonetheless incorrect (e.g., "rabbits" instead of *robins*). Two scenarios exemplified Type C errors.

Survey respondents were asked to provide brief narratives for up to three prompts for each of the six reading error scenarios. The prompts were categorised into word-level prompts (e.g., initial letter blends, letter-sound patterns), context-based cues (e.g., what makes sense in the story; look at the picture), and neutral prompts which included instructions by the teacher that did not relate specifically to any particular sources of information (e.g., "Are you sure?; "Keep trying"; "Have a go").

Literacy Teaching Efficacy Scale (LTES). The LTES was developed specifically for this project. Following the recommendations and guidelines for self-efficacy assessment (e.g., Bandura, 2006; Tschannen-Moran & Wolfok Hoy, 2001), a range of items was developed to assess teachers' beliefs about their capability of engaging in literacy teaching practices that would lead to desirable student learning outcomes. These items followed the stem, *I am confident I can…* Each item required the respondents to select their level of confidence on an 11-point scale, from "highly confident" (10) to "not at all confident" (0). For example, the first item was "*I am confident I can…*Create enthusiasm for reading among boys".

The LTES used in the present project was developed following piloting with a sample of 274 teachers spread throughout the country, but not in the geographical region of our research. A 30-item scale resulted from the pilot project. Cronbach's alpha was .98; the mean was 264.45 (SD = 44.50); the lowest score was 83 and the highest was 329, with the total possible score being 330. A principal components analysis of items revealed one strong factor that accounted for 61.96% of the variance.

APPENDIX 3

Details of the Teacher PLD Modules

Module 1: Introduction and the importance of language

In this module teachers were introduced to the cognitive development of reading framework, and the associated assessment framework. This first module included an introduction to effective instruction, including the roles of direct explicit instruction and implicit learning. This meant distinguishing between learning to read as learning to read 'sight words' and learning to 'work-out words.' The second part of module 1 was a deeper examination of the role of vocabulary in decoding and language comprehension, as well as an introduction to what phonological awareness is. 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 (Senechal, Ouellette, & Rodney, 2006; Tunmer & Chapman, 2012a, 2012b). Students 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 in text limits the ability to understand text. However, this does not suggest that instruction in foundation skills should be delayed until students have acquired fast, accurate word recognition skills (Tunmer & Chapman, 2012b).

Module 2: Understanding letter knowledge and phonological awareness: learning how to read words

In this module teachers were introduced to the specific developmental processes of letter knowledge and its relationship with phonological awareness, emphasising the way that they interact to contribute to alphabetic coding skills. A large body of scientific research indicates that comprehending text in an alphabetic orthography depends on the ability to recognize the words in 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). In this module teachers were provided with content knowledge distinguishing between vowels and consonants, how the sounds are similar and how they differ, as well as how students make use of sounding out for learning to read words independently.

Research on how students 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 automatisation 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).

Module 3: Developing word knowledge for fluency

In this module teachers were introduced to different word reading strategies that students need to learn, and how they are used in conjunction with each other. The teachers were provided with a scope-and-sequence, development progression, for the teaching of the different elements of phonic knowledge. They were also given specific instruction in the different long vowel sounds and digraphs, distinguishing between blend sounds and digraphs, and identifying morphemes in words. Another component of this module was the distinction between content knowledge (letters, sounds, phonic patterns, morphemes) and strategy instruction (how to make use of those components in reading as the way to read unfamiliar words).

Phonics instruction provides a 'kick-start' to phonological decoding for students who come to reading with few of the necessary cognitive entry skills, and who rely mostly on picture cues, partial visual cues, and sentence-context cues, with little interaction between the graphemes of printed words, and phonemes of spoken words (Tunmer & Greaney, 2010). For these students, the word recognition skills remain weak because they are unable to develop a rich network of sublexical connections between the orthographic and phonological representations in lexical memory. The use of inefficient word recognition processes drains the cognitive resources for comprehending the text being read.

Venezky (1999) argues that phonics instruction provides the processes by which learners can make estimates of the phonological representation of an unknown word. Explicit phonics instruction enables learners to explicitly produce approximate phonological representations (i.e., partial decodings) of unknown printed words (Tunmer & Arrow, 2013). These partial decodings are then used to generate alternative pronunciations of the words until one is found that matches a word in lexical memory and fits the context as well (Tunmer & Chapman, 2012a). The size of the reader's vocabulary is a critical component of the generation of alternative pronunciations. If a reader does not have the attempted word in their vocabulary they will not be able to come up with a suitable alternative and will be unable to induce the patterns from that word. When spelling-sound relationships are correctly identified they are stored with the accurate orthographic representation of words, which provide the data base from which further letter-sound patterns can be induced. Once students reach this point of development explicit instruction is not needed for word recognition and decoding.

Module 4: Reading comprehension as the goal

In this module teachers were introduced to direct instruction in comprehension instruction and how this can be introduced in junior classrooms. To reduce the negative Matthew effects in literacy there are three sources of variance that teachers must take into account: the reader, the text, and the activity engaged in (Snow, 2002). Connor and colleagues have found that attention to all of these aspects contributes to greater vocabulary development and reading comprehension outcomes in third grade classrooms (Connor, et al., 2014). This module looked at explicit reading comprehension strategy instruction and its place in the year 1 classroom and provided teachers with instruction in understanding the text (genre structure and how to teach it). However, as linguistic comprehension is necessary for reading comprehension it also looked at sentence construction and explicit instruction in sentences. Finally, in terms of Snow's (2002) notion of activity, teachers engaged in planning activities for each of the components of the Cognitive Framework covered (i.e., background knowledge in terms of genre and comprehension strategy use as well as sentence knowledge).

Initial comprehension instruction for beginning readers is less directed than word reading and vocabulary building. With beginning readers the pre-requisite abilities for

language comprehension, as indicated in Figure 1, are additional influences on reading comprehension (de Jong & van der Leij, 2002; Ouellette & Beers, 2010; Tunmer & Chapman, 2012b). These pre-requisites must first be assessed and identified before more dynamic comprehension-focused instruction can begin. Through the other aspects of reading already covered most students will be able to create basic meaning of text that is read; they will have sufficient cognitive processing abilities to do so as decoding skills become more efficient and a higher level of word level automaticity is achieved.

Module 5: Differentiated instruction as the goal

In this module teachers were focused on ideas for reconceptualising how to use whole class and small group instruction for the differentiated classroom from the start of the school year. Such changes had been introduced from Module 1 through the use of the templates guiding teachers to rethink their small group and whole class instruction including not only the how but also the what was taught in it. The long-standing approaches to reading in the junior classrooms are guided reading and shared reading (Ministry of Education, 2003). Vocabulary is critical so should be the cornerstone of instruction in beginning classrooms. Although the Language Experience approach is good at this, shared reading can build vocabulary beyond what Language Experience can do by the provision of text structure and vocabulary that students might not otherwise generate. Shared reading at the whole class level, during the first year of school, should emphasise the development of vocabulary and oral language, rather than as a means for introducing aspects of print and for developing fluency This approach encourages the use of multiple forms of shared book reading and reading aloud, including a variety of picture books (e.g., Braid, 2012) rather than the use of 'big books' alone.

Having an explicit knowledge of how students learn to read enables teachers to make informed instructional decisions that will move students forward. The use of specific assessments for beginning readers can also inform those decisions. Expectations are therefore based on what is known about the specific abilities, and what the next instructional steps should be. Another aspect of changing expectations is to be explicit in the use of direct instruction. This means telling students what they are learning and why they are learning it (Davis, 2007; Duffy, 2009). The small group instruction that beginning readers receive in the first year should not take the form of guided reading, in which students read their way through a text (Fountas & Pinnell, 1996; Ministry of Education,

2003). Rather, it should be rethought of as small group reading instruction that may include book reading. The focus, however, is on the explicit teaching of the specific abilities and skills that assessments have indicated many students need. The teaching, therefore, is planned based on need and not what arises from the text, as is currently the premise of guided reading.

APPENDIX 4a

Summary Beginning of Year 1 Student Assessment Data

	Intervention						Comparison					
	Low Decile n=47		Mid Decile n=106		High Decile n=34		Low Decile n=51		Mid Decile n=48		High Decile n=45	
Variables	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Letter Name	16.36	16.65	21.23	16.85	24.35	17.64	16.20	16.72	23.29	17.21	31.82	15.68
Letter Sound	7.40	14.19	10.32	14.22	13.06	15.63	9.24	15.17	13.63	14.75	18.29	15.76
CCTOP Elision	3.04	3.85	4.58	4.55	5.79	3.76	3.29	4.49	4.79	4.05	6.60	3.67
CTTOP Blends	6.60	3.48	6.74	4.59	6.91	4.57	5.71	4.01	5.94	4.13	7.51	4.11
CTTOP Matching	5.19	4.67	7.59	5.87	8.53	6.78	6.29	6.16	6.33	4.21	8.29	5.76
Clay words	0.36	0.97	0.40	1.15	0.97	2.54	0.51	2.22	0.67	2.19	0.62	1.19
Clay phonemes	1.40	3.84	1.12	2.77	2.71	6.91	1.59	6.01	2.19	6.15	2.09	3.88
Spelling	0.15	0.51	0.10	0.76	0.15	0.61	0.35	2.13	0.25	0.89	0.16	0.56
Spelling sounds	3.85	10.33	3.36	8.95	4.44	11.69	4.24	11.16	5.06	10.98	7.64	11.58
BPVS	92.81	10.50	100.29	11.84	99.85	10.69	93.41	11.29	97.67	10.91	104.98	10.39

APPENDIX 4b

Summary MANOVA/ANOVA Results for Beginning Year 1 Group Comparisons

A two-way (Group X Decile Band) Multivariate Analysis of Variance (MANOVA) on 10 Time 1 variables was performed to determine whether there were significant differences between the Intervention and Comparison groups. Results of the analyses are as follows:

Group Main Effect: $F(10,316)= 1.38$, $p=.19$. No significant difference between the Intervention and Comparison group.

Decile Band Main Effect: $F(20,634)=2.52$, $p<.001$. Statistically significant effect as a function of Decile Band. Univariate Analyses of Variance were run to identify for which variables the effects occurred:

- Letter Sound; $F(2,325) = 10.78$, $p = .000$.
- CTTOP Elision; $F(2,325) = 11.38$, $p = .000$.
- CTTOP Matching; $F(2,325) = 4.84$, $p = .001$.
- BPVS; $F(2,325) = 16.07$, $p = .000$.

Low decile students scored lower than high and middle decile students on these variables.

Group X Decile Band Interaction Effect: $F(20,634) = 1.27$, $p = .19$. Differences for students in the Intervention and Comparison groups in terms of Decile Band were not statistically significant.

APPENDIX 5

Details of Student Assessments Used at the End of Year 1

Phonological Processing. Phonological processing was assessed using the Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2: Wagner, Torgesen, Rashotte, & Pearson, 2013). Information about this test is reported in the earlier section on baseline assessments.

Alphabetic coding knowledge: To measure alphabetic coding we used four tasks that assessed phonic knowledge.

- The first was a test of 20 common consonant blends provided in lower case for students to provide the blended sound.
- The second was a test of the 5 most common digraphs (ph, sh, ch, wh, th) for which students were required to provide the accurate sound.
- The next measure assesses how well students apply phonic knowledge in their spelling attempts for the same spelling test used in Time 1. The words in this test are 18 real words that of the form of CVC (consonant-vowel-consonant) but may include a long vowel spelling or an initial or final consonant blend or digraph. In this scoring method for the 18 spelling words each word is given a score from 0-4. A score of 0 represents no attempt to spell the word or no match between what was written and the sound representation of the word, for example spelling '*gdv'* for CAKE. A score of 1 is given if one salient sound is accurately represented, a score of 2 is given if two sounds are accurately recorded and a score of 3 represents an accurate phonological representation but not conventional spelling. A score of 4 is given for accurate, conventional spellings. The maximum spelling score is 32.
- The pseudoword phoneme score comes from a pseudoword spelling test in which students are asked to read 30 made-up words of increasing complexity. The complexity comes from the word construction, from CVC words (jit) through to words containing consonant blends and both vowel and consonant digraphs (fleach). To assess alphabetic coding this was scored so that each accurate attempt to decode

a spelling pattern (consonant, short vowel, long vowel, digraph) was scored. From the 30 words a possible phoneme correct score is 101.

Language processing: Language processing measures how well a child can hold units of sound, or words in their working memory, as well as how the interaction between working memory and vocabulary works. Nonword repetition assesses the former and the mispronunciation task assesses the latter. These are not outcome tasks but can tell us a little about the influences on the outcomes.

- The nonword repetition task comes from the CTOPP test battery used for the phonological awareness assessments. In this task students are played a recording of a nonword and asked to repeat it. This continues until they are incorrect in their repeated pronunciation three times in a row. The maximum score is 34.
- In the mispronunciation task students are played a recording of an incorrectly pronounced word and asked what word the person in the recording meant to say. There are 40 items in this task of equal difficulty.

Reading and Spelling: Three outcome measures were used. Two tasks ask students to read or spell words and another, the pseudoword test measures how well students are able to use the content and strategies for decoding that were a key element of the workshop with teachers. Reading book level should also be included but we had difficulty in getting book levels for many students.

- The Burt word reading test is used as a standardised measure of students' word reading out of context. It is a good representation of the sight word reading knowledge of students rather than how well they can attempt to read unfamiliar words accurately.
- The pseudoword reading test is used as a reading outcome measure. Any word read correctly is scored as 1 and any incorrect words scored as 0. This provides a measure of how well students bring together both phonic knowledge and the strategy of sounding and blending unfamiliar words. This scoring system requires the blending of sounds in one fluently provided word rather than the correct sounding out without blending.

- The spelling score is taken from the spelling test described in the alphabetic coding section, but this score represents accurate conventional spelling; all correct words are scored as 1 and incorrect as 0.

Reading Book Level. Book level assessments are the most frequently used literacy assessments undertaken by New Zealand teachers. This was assessed at the end of Year 1 by the students' classroom teacher or other suitably qualified school personnel and provided to the project so was not independently assessed by the research assistants. Students are assigned to the book level in which they are able to attain a word recognition accuracy rate of 90-94%. Book level is not an equal interval scale as the average increase in book level for a given period of instruction is greater for the lower level books than for the higher level books. There are a total of 26 book levels, the characteristics of which are more fully described in Iversen and Tunmer (1993). Not all schools provided book level information as requested.

APPENDIX 6

Summary of End of Year 1 Student Assessments for Intervention Students as a Function of Teacher Change

	Same or other project teacher (n=234-265)		Change to non-project teacher or absent teacher (n= 34-36)		
Variables	M	SD	M	SD	t (df)
Blends	10.34	7.89	6.63	6.80	2.65**(294)
Digraphs	2.56	1.83	1.40	1.50	3.58**(294)
CTTOP Elision	12.55	6.51	9.37	6.25	2.73**(297)
CTTOP Blends	15.78	7.5	12.26	5.55	2.67**(298)
Pseudoword sounds	39.79	32.85	14.97	24.48	4.26**(290)
Invented spelling sounds	44.90	18.74	35.53	18.79	2.74**(295)
Burt word test	18.88	12.25	11.06	9.91	3.62**(297)
Reading Book Level	10.47	5.09	7.83	4.27	2.95**(268)

df = degrees of freedom
**$p < .01$

APPENDIX 7a

Summary End of Year 1 Results for Intervention and Comparison Students

	Intervention						Comparison					
	Low Decile n=28-34		Mid Decile n=50-73		High Decile n=29		Low Decile n=44-45		Mid Decile n=45-46		High Decile n=30	
Variables	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Blends	8.38	8.00	11.89	8.27	11.34	7.78	6.96	8.10	10.61	6.72	14.10	5.66
Digraphs	1.76	1.58	2.64	1.97	2.48	1.72	1.76	1.77	3.39	1.63	3.57	1.07
CCTOP Elision	11.24	6.25	12.53	7.37	13.21	4.69	11.53	6.96	12.85	6.31	15.87	4.09
CTTOP Blends	15.15	8.13	14.92	8.71	16.90	5.95	13.44	6.88	17.65	6.65	19.83	4.65
CTTOP Matching	14.00	5.59	13.82	7.55	16.52	5.88	13.82	7.17	16.37	6.06	18.73	6.00
Pseudoword snds	25.00	31.79	42.29	31.23	49.24	30.49	25.87	33.90	43.96	27.75	59.10	30.90
Invented sp snds	42.26	15.74	42.30	21.65	48.00	15.72	40.36	18.93	49.17	17.53	55.40	6.57
Burt word test	16.75	10.53	18.44	10.62	20.76	10.98	16.61	16.29	20.62	10.80	25.73	10.07
RDG Bk level	9.00	5.68	10.48	4.19	11.00	5.54	9.68	5.12	9.33	3.96	14.90	4.77
Spelling	3.96	3.06	4.44	3.41	5.28	3.47	4.11	3.84	6.09	3.51	6.83	2.88
Pseudoword rdg	3.04	4.37	4.10	5.54	5.38	5.82	3.57	5.86	6.33	6.24	9.77	8.04

APPENDIX 7b

Summary MANOVA/ANOVA Results for End of Year 1 Group Comparisons

Two-way (Group X Decile Band) Multivariate Analyses of Variance (MANOVA) on End of Year 1 variables were performed to determine whether there were significant differences between the Intervention and Comparison groups. Separate analyses were performed on "process" and "outcome" variables. Results of the analyses are as follows:

Process Variables

Group Main Effect: $F(7,246)=2.94$, $p<.01$. A statistically significant difference between the Intervention and Comparison groups. Univariate ANOVAs were run to identify the variables that contributed to the significant differences:

- Digraphs; $F(1,252)=7.04$, $p<.01$.

Decile Band Main Effect: $F(14,494)=3.75$, $p<.001$. Statistically significant effect as a function of Decile Band. Univariate Analyses of Variance were run to identify on what variables the effects occurred:

- Blends; $F(2,252) = 8.44$, $p = .001$.
- Digraphs; $F(2,252) = 15.02$, $p = .000$.
- Pseudoword phonemes $F(2,252) = 15.64$, $p = .000$.
- CTTOP Elision; $F(2,252) = 4.16$, $p < .05$.
- CTTOP Blending; $F(2,252) = 5.50$, $p < .01$.
- CTTOP Matching; $F(2,252) = 5.54$, $p < .01$.
- Spelling sounds; $F(2,252) = 5.59$, $p < 01$.

Low decile students scored lower than high and middle decile students on these variables.

Group X Decile Band Interaction Effect: $F(14,494) = 1.61$, $p = .07$. Students in the Intervention and Comparison groups performed at statistically similar levels in each of the three decile bands.

Outcome Variables

Group Main Effect: $F(4,217)=3.35$, $p < .01$. There was a statistically significant difference between the Intervention and Comparison groups. Results from ANOVAs showed the following significant differences:

- Pseudoword reading; $F(1,220) = 8.31$, $p < .01$.

- Spelling; $F(1,220) = 5.71$, $p = .02$.

Decile Band Main Effect: $F(8.436)=3.81$, $p<.001$. Statistically significant effect as a function of Decile Band. Univariate Analyses of Variance were run to identify on what variables the effects occurred:

- Burt Word Test $F(2,220) = 4.79$, $p < .01$.
- Reading Book Level $F(2,220) = 10.40$, $p < .001$.
- Pseudoword reading $F(2,220) = 7.91$, $p < .001$.
- Spelling $F(2,220) = 5.69$, $p < .01$.

High decile students in both groups tended to outperform low and middle decile students.

Group by Decile Band Interaction Effect: $F(8,436) = 5.03$, $p <.01$. This interaction effect was due to one significant univariate result.

- Reading Book Level $F(2,220) = 5.03$, $p < .01$. This effect was due to the high decile Comparison students outperforming the high decile Intervention students.

APPENDIX 8

Correlations Between Entry Variables and End of Year 1 Outcome Variables

Entry variables	Burt Word Test	Reading Book Level
Vocabulary knowledge	.42	.38
Letter Name	.66	.63
Letter Sound	.67	.61
Invented spelling sound	.58	.48
CTTOP Elision	.58	.56
CTTOP Blends	.44	.41
CTTOP Matching	.51	.49
Clay word phonemes	.52	.42

APPENDIX 9

Cohort 1 Students' Assessments at the End of Year 2

Foundations of literacy: By the end of Year 2 the foundations of literacy that still have some variability, were the Elision subtest of the CTOPP test and the pseudoword reading phoneme score. Both are already described.

Word Recognition: At the end of year 2, word reading and spelling was assessed with four tasks:

- The Burt word test as a measure of single word reading without context
- Connected text reading accuracy was measured with the Neale Analysis of Reading Ability (NARA: McKay & Barnard, 1999). In this measure students read passages of increasing length and difficulty until they make more than 16 errors per passage. The raw score was used in our analyses, using the raw score calculation from the NARA, which begins with a score of 16 per passage and errors per passage are deducted until the 16 errors in one passage threshold is reached. The maximum score over 6 passages is 100.
- The pseudoword reading words correct score was again used to measure decoding (implicit and explicit) without context or prior knowledge of a sight word to influence it.
- Spelling accuracy was measured with the Wide Range Achievement Test (WRAT-4) spelling test (Wilkinson & Robertson, 2008). This test has words of increasing difficulty with a maximum of 45 words. Students are given the words alone and in a sentence and asked to spell it until they get 10 consecutive words incorrect in a row. The raw score of words correct was used in our analyses.

Language comprehension: These assessments were used to measure language comprehension, which along with word recognition predict reading comprehension.

- Listening comprehension was measured using alternative passages from the NARA, which had been recorded in sound files using a native New Zealand speaker. Students were asked to listen to the passage and answer between 4-8 comprehension questions per passage. The test was stopped when the child gave 6

or more incorrect responses to the questions. The raw score of correct responses was used in our analyses, with a maximum score of 44.

- The mispronunciation task described for the end of Year 2 measures was again used.

Reading comprehension and fluency: Reading comprehension was the main outcome measure for the students at the end of Year 2. Reading fluency is included as it was measured alongside reading comprehension and can influence comprehension if a child reads too fast or too slow.

- The reading comprehension measure is the number of questions correctly answered after reading the NARA passages used for connected text reading accuracy. The passages had between 4 and 8 questions per passage that students were asked to answer. Any plausibly correct response to the question was acceptable. The maximum reading comprehension score was 44.
- Reading fluency was measured as the number of words read per minute when reading each NARA passage that had 16 or less accuracy errors within it. Across the 6 passages there were a total of 505 words that could be read.

APPENDIX 10a

Summary End of Year 2 Results for Intervention and Comparison Students

	Intervention						Comparison					
	Low Decile n=26-28		Mid Decile n=68-70		High Decile n=26-27		Low Decile n=34-38		Mid Decile n=47-48		High Decile n=28	
Variables	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
CCTOP Elision	15.19	5.34	17.07	6.74	17.56	6.62	16.29	7.63	19.00	5.74	20.21	5.61
Pseudoword snds	65.33	21.91	73.21	21.64	69.74	29.01	65.34	24.60	83.23	13.51	86.04	11.39
Burt word test	30.62	14.75	35.57	16.37	36.62	19.81	35.00	19.88	37.87	13.71	42.25	13.62
Reading accuracy	21.58	14.78	26.35	16.87	28.35	19.90	24.53	20.23	30.06	15.69	34.75	16.09
Pseudoword reading	8.12	7.47	11.96	8.45	11.65	8.48	9.50	8.54	15.62	8.91	17.43	8.24
WRAT spelling	4.50	2.93	5.49	2.72	5.42	2.93	5.62	3.98	6.06	2.57	6.14	3.02
Neale listening comp	5.75	3.90	7.39	5.29	7.96	4.86	6.70	4.77	8.54	4.95	9.89	5.35
Mispronunciation	10.81	5.64	12.29	7.27	13.15	7.94	11.03	6.76	17.94	9.35	13.64	5.79
Neale reading rate	45.78	31.68	50.17	21.71	48.96	26.51	46.19	29.62	54.98	21.73	61.18	20.20
Neale reading comp	6.19	4.03	8.63	5.61	10.22	6.91	7.79	6.94	9.13	5.11	12.46	4.40

APPENDIX 10b

Cohort 1 Summary MANOVA/ANOVA Results for End of Year 2

Several Group (2) by Decile Band (3) MANOVs/ANOVAs were performed on the following clusters of variables:

Foundations of literacy (CTOPP Elision; pseudoword reading phonemes).

Group Main Effect: $F(2,231) = 4.69$, $p < .01$. The Comparison group obtained higher scores than the Intervention group. Examination of univariate ANOVAs revealed the following significant effects:

- CTOPP Elision $F(1,232) = 4.59$, $p < .05$
- Pseudoword phonemes $F(1,232) = 9.38$, $p < .01$

Decile Band Main Effect: $F(4,464) = 4.56$, $p < .001$. Low decile students generally obtained lower scores than high and mid decile students. Both variables were statistically significant:

- CTOPP Elision $F(2,232) = 4.02$, $p < .05$
- Pseudoword phonemes $F(2,232) = 8.70$, $p < .001$

Group X Decile Band Interaction: $F(4,464) = 1.43$, $p = .22$. This effect was not statistically significant.

Word Recognition (Burt word test; Neale accuracy; pseudoword reading; WRAT spelling)

Group Main Effect: $F(4,220) = 2.69$, $p < .05$. The Comparison group generally obtained higher scores than the Intervention group. Examination of univariate ANOVAs revealed the following significant effect:

- Pseudoword reading $F(1,223) = 9.25$, $p < .01$

Decile Main Effect: $F(8,442) = 3.51$, $p < .01$. Low decile students generally obtained lower scores than high and mid decile students. Two variables were statistically significant:

- Neale accuracy $F(2,223) = 3.55$, $p < .05$
- Pseudoword reading $F(2,223) = 8.48$, $p < .001$

Group X Decile Interaction Effect: $F(8,442) = 1.18$, $p = .31$. This effect was not statistically significant.

Language Comprehension (Listening comprehension; Mispronunciation task)

Group Main Effect: $F(2,230) = 3.59$, $p < .05$. Comparison students generally performed better than Intervention students on both variables:

- Listening comprehension $F(1,231) = 3.88$, $p = .05$
- Mispronunciation task $F(1,231) = 5.18$, $p < .05$

Decile Band Main Effect: $F(4,462) = 5.26$, $p < .001$. Low decile students generally obtained lower scores than high and mid decile students. Both variables were statistically significant:

- Listening comprehension $F(2,231) = 4.67$, $p < .01$
- Mispronunciation task $F(2,231) = 7.52$, $p < .001$

Group X Decile Band interaction effect: $F(4,462) = 1.88$, $p = .11$. The interaction effect was not statistically significant.

Reading comprehension and fluency (Neale reading comprehension; Neale fluency)

Group Main Effect: $F(2,224) = 1.99$, $p = .14$. No significant difference between the Intervention and Comparison groups.

Decile Band Main Effect: $F(4,450) = 4.51$, $p < .001$. Low decile students generally scored lower that high and mid decile students. One variable was statistically significant:

- Reading comprehension $F(2,225) = 8.68$, $p < .001$.

APPENDIX 11

Cohort 1 Middle of Year 3 Assessments

The following Middle of Year 3 assessments have not been described elsewhere in this report.

- **Morpheme test**. This test involved a word analogy task that required the student to first listen to a pair of related words. When the third word is read, the student was required to repeat the relationship found between the first pair; e.g., *tall*; *tallest*, *small*; **smallest**. Patterns included changes in tense, singular to plural and related verbs and nouns. The assessment contained fourteen sets, with a point scored for each correct answer (Tong, Deacon & Cain, 2014).
- **Word attack strategy**. This task involved simply asking students the following open-ended question: "When you are reading on your own and come across a word you don't know, what do you do to figure out what the word is?" Responses were scored in one of two categories: word-related word attack strategy, such as sounding out, mentioning the letters; or non-word strategy, such as guessing, looking at the picture, asking the teacher, other.
- **Reading self-efficacy** as an indicator of motivation, was assessed with a Reading Self-Efficacy Scale. Six reading self-perception statements were presented to students (e.g., "I usually do well in reading"; "Reading is easy for me"; "I have trouble reading stories with difficult words"). They were asked to rate their agreement with each statement on a 4-point scale ranging from "Agree A Lot" to "Disagree A Lot". An equal number of positive and negative statements was included in the scale, and scoring was reversed where appropriate so that"4" always represented a more positive self-perception and "1" was indicative of a negative self-perception.

APPENDIX 12a

Summary Middle of Year 3 Results for Intervention and Comparison Students

	Intervention						Comparison					
	Low Decile n=23-27		Mid Decile n=68-69		High Decile n=19-24		Low Decile n=36-38		Mid Decile n=43-44		High Decile n=28-29	
Variables	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Burt Word Test	36.33	15.77	41.29	17.76	42.71	19.45	34.00	16.10	43.98	15.67	49.41	15.48
Reading Accuracy	58.42	29.88	60.28	24.11	60.44	24.61	57.17	26.66	59.28	22.00	76.30	20.72
WRAT Spelling	5.81	3.74	7.22	4.00	6.92	4.26	5.49	3.68	7.16	2.96	8.03	3.74
Neale Listening Comp	7.04	3.41	8.98	5.64	11.88	6.20	6.89	4.02	9.81	5.28	12.96	6.10
Morpheme Test	5.22	3.50	6.86	4.05	6.96	2.71	5.81	3.37	7.26	4.40	9.04	4.23
Mispronunciation	14.30	6.79	14.98	7.28	14.63	6.47	11.81	5.96	16.24	6.36	17.81	7.09
BPVS Standard Score	89.70	10.22	95.32	11.75	98.21	11.00	88.95	10.49	93.79	9.62	99.61	8.04
Reading Book Level	19.74	3.41	20.36	4.19	19.05	3.47	16.28	6.08	20.79	4.52	21.59	1.59
Reading Self-Efficacy	18.12	2.79	18.04	2.82	18.17	2.65	18.13	2.59	18.07	3.23	19.59	2.41

APPENDIX 12b

Cohort 1 Summary MANOVA/ANOVA Results for Middle of Year 3

Word Recognition (Burt Word Test; Neale Accuracy; WRAT Spelling)

Group Main Effect: $F(3,221) = 0.79$, $p = .50$. Not significant.

Decile Main Effect: $F(6,444) = 2.99$, $p < .01$. All three variables were statistically significant, with low decile students generally performing less well than high and middle decile students.

- Burt word: $F(2,223) = 6.61$, $p < .01$
- Accuracy: $F(2,223) = 3.01$, $p = .05$
- Spelling: $F(2,223) = 4.40$, $p = .01$

Group X Decile Band interaction effect: $F(6,444) = 1.19$, $p = .31$. The interaction effect was not statistically significant.

Language Comprehension (Listening comprehension; morpheme; mispronunciation; BPVS)

Group Main Effect: $F(4,212) = 1.55$, $p = .19$. Not statistically significant.

Decile Main Effect: $F(8,426) = 4.56$, $p < .001$. All four variables were statistically significant, with low decile students tending to perform lower than high and mid decile students.

- Listening comprehension $F(2,215) = 15.69$, $p < .001$
- Morpheme $F(2,215) = 6.18$, $p < .01$
- Mispronunciation $F(2,215) = 3.48$, $p < .05$
- BPVS $F(2,215) = 12.04$, $p < .001$

Group by Decile Interaction Effect: $F(8,426) = 1.00$, $p = .44$. Not statistically significant.

Reading Comprehension & Fluency (Neale Reading Comprehension; Neale Fluency)

Group Main Effect: $F(2,222) = 0.89$, $p = .41$. Not statistically significant.

Decile Main Effect: $F(4,446) = 7.16$, $p < .001$. Both variables were statistically significant, with low decile students tending to perform lower than high and mid decile students.

- Reading comprehension $F(2,223) = 14.00$, $p < .001$

- Fluency $F(2,223) = 3.01$, $p = .05$

Group by Decile Interaction Effect: $F(4,446) = 1.32$, $p = .26$. Not statistically significant.

Reading Book Level (Univariate ANOVA)

Group Main Effect: $F(1,213) = 0.07$, $p = .79$. Not statistically significant.

Decile Main Effect: $F(2,213) = 7.01$, $p < .001$. Low decile students lower than high and mid decile students.

Group by Decile Interaction Effect: $F(2,213) = 6.71$, $p < .001$. This effect was due to the Intervention low decile group performing significantly higher than the Comparison low decile group.

Reading Self-Efficacy (Univariate ANOVA)

Group Main Effect: $F(1,224) = 1.53$, $p = .22$. Not statistically significant.

Decile Main Effect: $F(2,224) = 1.63$, $p = .20$. Not statistically significant.

Group by Decile Interaction Effect: $F(2,224) = 1.25$, $p = .29$. Not statistically significant.

APPENDIX 13a

Cohort 2 Means and standard deviations for Year 1 School Entry Assessments.

	Intervention						Intervention+						Comparison					
	Low Decile		Middle Decile		High Decile		Low Decile		Middle Decile		High Decile		Low Decile		Middle Decile		High Decile	
	N=43-44		N=42		N=30-31		N=33-35		N=90-91		N=9		N=12-13		N=21		N=24	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Phono Awareness																		
Elision	3.34	4.44	2.98	2.84	5.23	3.79	3.46	6.00	4.30	3.43	5.11	3.66	2.92	3.42	4.29	3.24	5.92	3.19
Matching	4.46	4.87	5.45	4.68	9.43	5.38	5.71	5.37	5.78	4.69	5.78	4.15	4.85	6.04	5.00	3.86	7.75	5.28
Blending	7.09	4.08	5.24	3.67	6.57	2.50	5.71	3.88	5.82	3.38	4.56	4.22	5.69	3.22	6.33	2.90	7.25	4.10
Letter Identification																		
Letter Sounds	4.89	13.12	9.86	11.90	11.30	16.97	8.89	14.65	14.24	15.39	13.56	16.95	3.15	5.76	8.90	9.96	20.44	17.74
Letter Names	13.91	17.73	16.95	16.14	23.00	18.51	14.46	15.38	22.88	17.85	22.00	19.68	6.92	9.35	20.86	14.19	25.76	19.32
Receptive Vocabulary																		
BPVS	93.63	11.57	100.50	10.97	105.19	9.16	91.33	10.15	100.91	10.35	103.33	12.98	91.67	11.48	97.57	11.76	96.67	10.31
Reading & Spelling																		
Clay word test	0.39	1.73	0.33	1.56	0.57	0.97	0.29	0.79	0.62	1.52	0.56	0.73	0.08	0.29	0.10	0.44	1.25	1.48
Invented Spelling	0.02	0.15	0.02	0.15	0.20	0.66	0.09	0.51	0.08	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.13

APPENDIX 13b

Cohort 2 Summary MANOVA/ANOVA Results for Beginning of Year 1

Phonological Awareness (CTOPP Elision, blending, matching)

Group Main Effect: $F(6,598) = 1.28$, $p = .27$. Not significant.

Decile Main Effect: $F(6,598) = 2.61$, $p = .02$. Two variables were statistically significant. In each case, high decile students in both groups had higher scores than low and middle decile students.

- Elision: $F(2,300) = 4.61$, $p = .01$
- Matching: $F(2,300) = 4.58$, $p = .01$

Group X Decile Band interaction effect: $F(12,900) = 1.22$, $p = .27$. Not significant.

Letter Identification (Letter Names, Letter Sounds)

Group Main Effect: $F(4,602) = 1.11$, $p = .35$. Not significant.

Decile Main Effect: $F(4,602) = 4.37$, $p < .01$. Both Letter Name and Letter Sound were statistically significant.

- Letter Name $F(2,301) = 8.16$, $p < .001$
- Letter Sound $F(2,301) = 6.50$, $p < .01$

Low decile students scored lower than middle and high decile students in both groups on both variables.

Group X Decile Band interaction effect: $F(8,602) = 1.94$, $p = .05$. An examination of the univariate ANOVAs revealed that neither of the two variables had a statistically significant interaction effect.

Reading and Spelling (Clay Word Test, Invented Spelling)

Group Main Effect: $F(4,596) = 0.42$, $p = .80$. Not significant.

Decile Main Effect: $F(4,596) = 1.27$, $p = .28$. Not significant.

Group X Decile Band interaction effect: $F(8,596) = 1.39$, $p = .20$. Not significant.

Receptive Vocabulary (BPVS)

Group Main Effect: $F(2,296) = 3.12$, $p = .05$. There was a tendency for the Comparison students to obtain lower scores than the Intervention groups, but post hoc individual comparisons of means did not reveal statistically significant differences between the Comparison group and either the Intervention or Intervention^{+} groups.

Decile Main Effect: $F(2,296) = 13.89$, $p = <.01$. Students in the low decile group obtained lower scores than students in the middle and high decile groups, across all the samples.

Group X Decile Band interaction effect: $F(4,296) = 0.91$, $p = .46$. Not significant.

APPENDIX 14a

Cohort 2 Means and Standard Deviations for End of Year 1 Assessments

	Intervention						Intervention+						Comparison					
	Low Decile N=43-44		Middle Decile N=42		High Decile N=30-31		Low Decile N=33-35		Middle Decile N=90-91		High Decile N=9		Low Decile N=12-13		Middle Decile N=21		High Decile N=24	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Phono Awareness																		
Elision	11.27	6.00	11.08	6.60	14.52	3.62	9.44	5.96	12.64	5.27	9.50	5.66	4.47	5.01	12.00	4.93	12.52	6.08
Matching	14.88	7.46	16.35	6.98	18.35	4.98	14.37	6.93	15.34	6.71	13.10	4.82	7.73	4.48	13.48	6.62	16.56	7.59
Blending	14.85	6.03	16.14	6.77	19.19	4.45	11.34	6.56	15.94	5.43	14.30	5.06	7.40	3.81	14.33	5.35	15.84	6.48
Alphabetic Coding																		
Blends	11.00	8.52	10.54	7.28	16.62	4.20	8.04	7.56	10.38	7.26	6.67	7.97	0.75	1.39	9.10	8.04	12.38	7.96
Digraphs	2.36	1.90	3.57	1.38	3.14	1.68	2.08	1.69	2.44	1.73	1.56	2.01	0.25	0.71	2.00	1.97	3.17	1.81
Spelling phonemes	45.36	15.81	52.40	14.71	52.21	10.31	42.42	18.42	46.78	16.47	43.78	15.25	17.13	15.81	38.45	18.86	46.37	17.31
Language Processing																		
Mispronunciation	8.68	4.94	12.37	9.39	9.45	4.02	7.25	4.80	6.94	4.10	5.70	2.79	3.93	2.25	6.67	3.28	6.80	3.38
Nonword rep	13.89	6.25	12.37	3.61	17.03	3.39	15.13	4.00	13.95	3.69	10.70	5.96	10.33	5.02	14.86	2.76	14.88	3.93
Reading & Spelling																		
Burt word test	18.20	12.08	20.85	11.69	21.26	9.96	14.11	11.18	19.41	10.34	18.20	10.60	4.00	3.85	14.24	7.77	21.25	9.04
Pseudoword rdg	5.06	6.141	7.47	6.23	7.94	6.78	4.00	5.72	4.72	5.33	5.70	5.52	0.50	1.41	3.86	6.39	5.65	4.84
Spelling	4.37	2.88	6.62	3.61	5.29	2.48	3.56	3.12	4.74	3.31	4.00	2.91	0.50	1.07	3.29	2.70	4.65	2.79

APPENDIX 14b

Cohort 2 Summary MANOVA/ANOVA Results for End of Year 1

Phonological Awareness (CTOPP Elision, blending, matching)

Group Main Effect: $F(6,580) = 4.00$, $p < .01$. All three variables showed statistically significant differences.

- Elision: $F(2,291) = 4.77$, $p < .01$ (Intervention higher than Comp)
- Blending: $F(2,291) = 11.40$, $p < .001$ (Intervention higher than I^+ & Comp)
- Matching: $F(2,291) = 7.02$, $p < .01$ (Intervention higher than Comp)

Decile Main Effect: $F(6,580) = 5.48$, $p < .001$. All three variables showed statistically significant differences, with the Low Decile group scoring significantly lower than the middle and high Decile groups for each variable.

- Elision: $F(2,291) = 10.65$, $p < .001$
- Blending: $F(2,291) = 16.45$, $p < .001$
- Matching: $F(2,291) = 5.60$, $p < .01$

Group X Decile Band interaction effect: $F(12,873) = 2.14$, $p = .01$. This interaction effect was due to the Low Decile Intervention group scoring higher than the Low Decile Comparison group.

- Elision: $F(4,291) = 4.66$, $p < .01$
- Blending: $F(4,291) = 2.49$, $p < .05$
- Matching: $F(4,291) = 2.51$, $p < .05$

Alphabetic Coding (Blends, Digraphs, Spelling phonemes)

Group Main Effect: $F(6,510) = 6.47$, $p < .01$. All three variables showed statistically significant differences.

- Blends: $F(2,256) = 10.72$, $p < .001$ (Intervention higher than I^+, Comp)
- Digraphs: $F(2,256) = 10.05$, $p < .001$ (Intervention higher than I^+, Comp)
- Spelling phonemes: $F(2,256) = 14.70$, $p < .001$ (Intervention higher than Comp)

Decile Main Effect: $F(6,510) = 4.11$, $p < .001$. All three variables showed statistically significant differences, with the Low Decile group scoring significantly lower than the middle and high Decile groups for each variable.

- Blends: $F(2,256) = 6.48$, $p < .01$ (Low decile < high decile)
- Digraphs: $F(2,256) = 7.07$, $p < .01$ (Low decile < middle, high decile)
- Spelling phonemes: $F(2,256) = 9.12$, $p < .001$ (Low decile < middle, high decile)

Group X Decile Band interaction effect: $F(12,768) = 5.60$, $p < .001$. This interaction effect was due mainly to the Low Decile Intervention group scoring higher than the Low Decile Comparison group.

- Blends: $F(4,256) = 3.91$, $p < .01$
- Digraphs: $F(4,256) = 3.94$, $p < .01$
- Spelling phonemes: $F(4,256) = 2.84$, $p < .05$

Language Processing

Group Main Effect: $F(4,572) = 9.34$, $p < .001$. There was a statistically significant effect for Mispronunciation, but not for Nonword repetition.

- Mispronunciation $F(2,286) = 19.24$, $p < .001$. The Intervention group obtained statistically significantly higher scores than the Intervention^{+} and Comparison groups.

Decile Band Main Effect: $F(4,572) = 2.48$, $p < .05$. The univariate ANOVA showed that this effect was due to the Mispronunciation variable.

- Mispronunciation $F(2,286) = 4.05$, $p < .05$. None of the univariate ANOVAs were statistically significant.

Group by Decile Band Interaction Effect: $F(8,572) = 5.83$, $p < .001$. Both variables were statistically significant.

- Mispronunciation $F(4,286) = 2.54$, $p < .05$. The low decile Intervention group obtained higher scores than the low decile Comparison group.
- Nonword repetition $F(4,286) = 8.56$, $p < .001$. The low decile Intervention group obtained higher scores than the low decile Comparison group. The high decile Intervention+ group obtained lower scores than the other high decile groups.

Reading & Spelling

Group Main Effect: $F(6,522) = 3.87$, $p < .001$. All three variables had statistically significant group differences.

- Burt word test $F(2,262) = 6.58$, $p < .01$. Intervention students had higher scores than Comparison students.
- Pseudoword reading $F(2,262) = 5.90$, $p < .01$. Intervention students had higher scores than Intervention^{+} and Comparison students.
- Spelling $F(2,262) = 11.50$, $p < .001$. The intervention group obtained higher scores than the Intervention^{+} and Comparison groups.

Decile Band Main Effect: $F(6,522) = 4.03$, $p < .001$.All three variables had statistically significant Decile Band effects.

- Burt word test $F(2,262) = 8.29$, $p < .001$ The low decile group obtained lower scores than the middle and high decile groups.
- Pseudoword reading $F(2,262) = 4.00$, $p < .05$. The low decile group obtained lower scores than the high decile group, but not the middle decile group.
- Spelling $F(2,262) = 7.81$, $p < .001$. Students in the low decile group generally obtained lower scores than students in the middle and high decile groups.

Group by Decile Band Interaction Effect: $F(12,786) = 2.11$, $p < .05$. Only the Burt word test resulted in a statistically significant effect.

- Burt word test $F(4,262) = 2.43$, $p < .05$. The low decile Intervention group obtained significantly higher scores than the low decile Comparison group.

APPENDIX 15a

Cohort 2 Middle of Year 2 Assessment Means and Standard Deviations

	Intervention						Intervention+						Comparison					
	Low Decile N=32-33		Middle Decile N=32-34		High Decile N=29-30		Low Decile N=21-22		Middle Decile N=84-87		High Decile N=9		Low Decile N=10		Middle Decile N=10-17		High Decile N=18-24	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Language Compr																		
Listening comp	5.12	3.79	6.81	4.99	8.00	6.01	3.90	2.66	6.07	3.81	8.78	4.74	3.10	2.47	4.90	2.38	6.92	5.17
Nonword rep	4.97	4.87	16.62	3.11	18.59	2.77	14.57	5.00	14.99	3.47	15.56	3.43	15.50	2.72	13.00	4.57	14.00	6.42
Morpheme	3.63	2.97	4.87	3.46	5.21	2.91	2.29	2.39	4.92	3.10	4.67	3.39	3.10	1.66	4.20	2.66	4.29	2.51
Receptive vocab	90.59	11.04	96.75	11.28	95.79	10.41	87.38	6.35	95.70	9.14	98.78	9.35	88.10	8.16	96.30	8.15	91.38	8.41
Reading & Spelling																		
Burt word test	27.09	12.22	29.15	13.32	30.67	11.26	21.45	13.06	29.03	14.73	24.22	14.63	12.73	12.76	23.65	8.41	29.48	13.72
Pseudoword rdg	9.30	8.02	9.38	9.03	13.07	7.53	5.03	6.18	8.06	6.99	7.89	8.04	4.45	5.89	4.35	4.12	8.17	6.31
WRAT Spelling	4.27	2.59	4.74	3.67	4.57	2.46	2.81	1.80	4.44	4.02	3.89	2.26	1.82	1.17	3.00	1.41	4.52	2.25
Reading Book Level	13.37	5.01	14.76	4.86	15.63	3.98	12.77	6.81	15.60	4.48	13.00	5.05	5.50	2.84	12.47	3.54	14.89	5.50
Reading Self-Efficacy	17.94	3.42	17.71	4.03	17.77	4.07	17.47	4.38	17.00	3.85	17.00	4.61	18.50	2.15	16.94	2.79	15.50	3.36

APPENDIX 15b

Cohort 2 Summary of Analyses for Middle of Year 2 Assessments

Language Comprehension (Listening comprehension, Nonword repetition, Receptive vocabulary, Morpheme knowledge).

Group Main Effect: $F(8,480) = 1.86$, $p = .06$. Marginally significant. One variable was statistically significant.

- Nonword Repetition $F(2,242) = 6.23$, $p < .01$ Intervention group higher scores than Intervention$^+$ and Comparison groups.

Decile Band Main Effect: $F(8,480) = 4.50$, $p < .001$. Three of the four variables were statistically significant. In each case, the low decile group obtained lower scores than the middle and high decile groups.

- Listening comprehension $F(2,242) = 9.87$, $p < .001$
- Receptive vocabulary $F(2,242) = 9.86$, $p < .001$
- Morpheme knowledge $F(2,242) = 5.62$, $p < .001$

Group by Decile Band Interaction Effect: $F(16,968) = 1.05$, $p = .41$. Not statistically significant.

Reading and Spelling (Burt word test, pseudoword reading, WRAT spelling)

Group Main Effect: $F(6,530) = 3.13$, $p < .01$. Statistically significant differences occurred for all three variables in this cluster.

- Burt word test $F(2,266) = 4.78$, $p < .01$ Intervention higher scores than Comparison group.
- Pseudoword reading $F(2,266) = 8.95$, $p < .001$. Intervention group higher scores than Intervention$^+$ and Comparison groups.
- WRAT spelling $F(2,266) = 3.59$, $p < .05$. Intervention group higher scores than Comparison group.

Decile Band Main Effect: $F(6,530) = 4.40$, $p < .001$. All three variables had statistically significant differences. The low decile group obtained lower scores on all variables than the high decile group, and lower scores than the middle decile group for the Burt word test and Spelling.

- Burt word test $F(2,266) = 6.05$, $p < .01$

- Pseudoword reading $F(2,266) = 3.17$, $p < .05$
- WRAT spelling $F(2,266) = 3.17$, $p < .05$

Group by Decile Band Interaction Effect: $F(12,798) = 1.99$, $p < .02$. The low decile Intervention group obtained higher scores than the low decile Comparison group, however, none of the univariate analyses was statistically significant.

APPENDIX 16a

Reading and Spelling Means and Standard Deviations for Cohort 1 and Cohort 2 Intervention Groups

	Cohort 1 Intervention						Cohort 2 Intervention					
	Low Decile n=36		Mid Decile n=96		High Decile n=30		Low Decile n=33		Mid Decile n=34		High Decile n=30	
Variables	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Burt word test	22.89	13.69	26.14	16.85	30.20	18.88	27.09	12.22	29.15	13.33	30.67	11.26
Pseudoword reading	6.19	6.27	7.77	8.15	9.83	7.97	9.30	8.02	9.38	9.03	13.07	7.53
WRAT spelling	3.03	1.78	3.78	2.63	4.17	2.36	4.27	2.59	4.74	3.67	4.57	2.46

APPENDIX 16b

Summary Reading and Spelling Analyses Data Comparing Cohorts 1 and 2 Intervention Groups.

Reading and Spelling (Burt word test, pseudoword reading, WRAT spelling).

Group Main Effect: $F(3,251) = 3.30$, $p = .02$. Cohort 2 Intervention students outperformed Cohort 1 Intervention students for Pseudoword reading and WRAT spelling. The difference in Burt word scores was not statistically significant.

- Pseudoword reading $F(1,253) = 6.10$, $p = .01$
- WRAT spelling $F(1,253) = 5.89$, $p = .02$

Decile Band Main Effect: $F(6,504) = 2.50$, $p = .02$. The univariate analyses revealed one statistically significant effect.

- Pseudoword reading $F(2,253) = 3.85$, $p = .02$ The high decile group obtained significantly higher scores than the middle and low decile groups.

Group by Decile Band Interaction Effect: $F(6,504) = 0.96$, $p = .45$. Not statistically significant.

APPENDIX 17

Teacher Survey Time 1

Teacher Knowledge. As for Cohort 1

Word Identification Prompt Scenarios. The word identification prompt task was based on scenarios developed by Greaney (2001) and adapted to suit an examination of teaching reading in the first year of school. These scenarios were selected from the published books most commonly used in New Zealand junior classes, the *Ready to Read* series. The scenarios were created to exemplify common error types for beginning readers. The errors were an incomplete attempt, where the reader used the initial letter only (l---- for 'leader'); an omission error where the child missed a word; two examples of a substitution error where a replacement word was read ('get' for 'take'; 'weared' for 'wore') and a non-word replacement (crooms for crumbs).

Survey respondents were asked to provide brief narratives for up to three prompts for each of the six reading error scenarios. The prompts were categorised into word-level prompts (e.g., initial letter blends, letter-sound patterns), context-based cues (e.g., what makes sense in the story; look at the picture), and neutral prompts which included instructions by the teacher that did not relate specifically to any particular sources of information (e.g., "Are you sure?"; "Keep trying"; "Have a go"). The teachers' prompts were then analysed on the basis of whether a word level prompt was used as a first prompt and whether the prompt included a focus on a decoding strategy, both of which were features of the intervention workshops.

Table 5 *Scenarios for teacher prompts*

Text used (Ready to Read)	Reading error	Error type
1. Magenta level 1-2 *The Waterslide*	Dad turned the hose. Dad turned on the hose.	Omission
2. Red level 3-5 *A starfish for Oscar*	We can get a photo We can take a photo	Substitution
3. Yellow level 6-8 *A Bird in the classroom*	We put some crooms by the door We put some crumbs by the door	Medial vowel
4. Blue level 9-11 *Stay where you are*	I will be the l . I will be the leader.	Incomplete attempt
5. Green level 12-14 *Dad's hat*	He weared the hat everywhere He wore the hat everywhere	Over-generalised tense

Rubric for structured observation. The structured observation measure was developed specifically for this Project. The first step in developing the rubric was to examine how other studies had developed such scales (Chen, Hu, Fan, & Li, 2014; Connor, 2013; Crawford, Zucker, Williams, Bhavsar, & Landry, 2013; Doabler & Nelson-Walker, 2009); Pianta, La Paro, & Hamre, 2008; Reddy, & Dudek, 2014; Reddy, Fabiano, & Peters, 2015; Walpole & McKenna, 2013). A number of trials of indicators and rating schemes resulted in a rubric of six elements of small group instructional practice. The six elements were lesson focus, teaching strategies, teaching code knowledge, text selection, reader strategies, and materials used. The rubric enabled a rating of each of the elements as either incidental, implicit, explicit, or systematic with indicators to describe practice for each rating. The indicators in the explicit and systematic categories were developed from the content of the intervention workshops, while the indicators for incidental and implicit were developed from viewing teacher current practice and from commonly used literacy materials.

APPENDIX 18

Results for basic linguistic constructs test at time 1 and time 2

	Max score	Time 1 Mean (SD)	Time 2 Mean (SD)	*t*	*p*	*D*
Total knowledge	38	23.19 (5.79)	28.57 (4.82)	5.35	.00	1.00
Explicit knowledge	12	5.33 (2.61)	8.05 (2.29)	5.11	.00	1.11
Implicit knowledge	26	17.95 (3.70)	20.48 (3.50)	3.67	.00	0.70
Phonemic	13	9.33 (2.52)	10.33 (2.00)	3.02	.00	0.50
Phonological	8	7.14 (0.57)	7.55 (0.75)	2.26	.03	0.61
Phonic	9	4.24 (1.84)	6.13 (1.67)	4.52	.00	1.07
Morphological	8	2.55 (2.68)	4.81 (2.41)	3.12	.00	0.89

APPENDIX 19

Teachers' self-evaluation ratings for each aspect of the survey

	T1	T2			
	Mean (SD)	Mean (SD)	*t*	*p*	*d*
Phonemic	2.16 (0.69)	2.68 (0.48)	3.29	.00	0.87
Phonics	2.37 (0.60)	2.79 (0.54)	3.62	.00	0.81
Fluency	2.37 (0.68)	2.63 (0.50)	1.76	.10	0.58
Vocabulary	2.42 (0.61)	2.79 (0.54)	3.24	.01	0.67
Comprehension	2.58 (0.61)	2.63 (0.50)	.38	.17	0.17
Children's Literature	2.53 (0.51)	2.58 (0.51)	.57	.10	0.10
English Language Learners	2.00 (0.76)	2.37 (0.60)	2.69	.02	0.59
Reading Assessment	2.53 (0.70)	2.84 (0.50)	2.36	.03	0.58

APPENDIX 20

Small group reading instruction for Year 1 students	1: Incidental	2: Implicit	3: Explicit	4: Systematic
1. Lesson focus	Reading the levelled book is lesson focus. Purpose to find out what happens in story and discuss key ideas. Focus on fluent reading and meaning.	Book is lesson focus. Purpose is to read for meaning and to practice reading new words. Fluent reading as follow up focus	Explicit teaching for knowledge and the words and sentences Ss need for a successful reading of the selected text. Focus on accuracy.	Explicit teaching of knowledge and strategies based on learner need. Systematically follows scope and sequence. Focus on using strategies
2. Instructional strategies	T Qs as main strategy, leading Ss to self-monitor. "*Did that sound right*?"; "*What did…*?"; "*What's this letter/sound*?"	T Qs with prompts added when Ss need support e.g. "*Look at…*" "*Point to…*" "*Let's try…*" "*Make the sound*"	T uses prompts to guide students in new learning and explains or models when explicit teaching needed.	T explains and models new learning, then prompts or questions towards independent use of strategies.
3. Code teaching	Ss learn code while reading connected text (a book). T supports decoding attempts as need arises in text.	Alphabet or HF words reviewed prior to reading text. Decoding support as need arises in the text.	Explicit teaching for letters and words for book. HF words taught as whole. Letter formation practiced. Decoding strategies modelled.	Teaching code knowledge[2] and decoding strategies[3] as priorities at early levels. HF words and letter formation taught and practiced.
4. Choice of text	Text has natural language[4] and strong storyline selected for Ss instructional level. Text includes words outside Ss known reading vocabulary	Controlled language[5], repetitions of HF words and sentence patterns to support learners. Text is chosen for Ss instructional level.	Text carefully selected to support the explicit teaching focus in knowledge. Controlled or natural language selected as appropriate.	Text selected to support explicit teaching of knowledge and strategy. Decodable texts or teacher made sentences related to phases.
5. Reader strategies	T directs Ss to problem solve using processing system[6], m/s and first letter to predict.	T supports Ss to use processing system, m/s with word parts to predict.	T supports Ss to use code knowledge with m/s to predict or confirm.	Ss apply knowledge and strategies taught in lesson. T guides Ss to use code first with m/s to confirm.
6. Materials used	Instructional book is the material in the lesson.	Instructional book; modelling resource. T uses	Text, modelling resource, magnetic letters, word/letter cards. T uses.	Text, modelling resource, magnetic letters, word/letter cards, whiteboards. T and Ss use.

T=Teacher Ss = students HF: high frequency Q = questions m=meaning s=structure

[2] **Code knowledge:** learning the letter sounds, blends, vowel patterns, syllables and morphemes as appropriate to learner phase

[3] **Decoding strategies**: sounding/ blending, using small then bigger chunks, recognizing and using vowel patterns, using syllables and morphemes as appropriate to learner phase

[4] **Natural language**: as close to natural speech and sentence structure as possible; full use of punctuation range (Ready to Read).

[5] **Controlled:** based on speech patterns but with careful control of vocabulary in a levelled progression; full use of punctuation range (PM plus);

[6] **Processing system:** self-monitor, re-run, check 1:1, read on, use 3 cues of meaning (picture), sentence structure, and as much of the printed code as needed