Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.
http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
The use of evidence in reporting on reading: A study into how overall teacher judgements are made when assessing students’ reading in relation to the National Standards.

Joanne Ruth Quick

Abstract

National Standards in reading require that teachers make an overall teacher judgement (OTJ) of students’ reading by considering multiple sources of assessment evidence in relation to the Standards. The lack of recent professional learning and development (PLD) in teaching and assessing reading, together with teachers’ requests to the researcher for support in this area, indicated that making reading OTJs was challenging for teachers. Early studies of teachers’ work with the Standards reported mixed experiences with PLD and in making OTJs. Consistent trends identified were that schools struggled to align existing assessment practices with the Standards, and to gain consistency of OTJs. There was little data specifically relating to reading OTJs. A small scale mixed methods study was conducted across a semi-rural cluster of primary schools to investigate the tools and processes teachers of different age groups used to make OTJs in reading, to investigate the usefulness of PLD accessed, to probe teachers’ levels of confidence in their OTJs and to inquire how schools are working towards consistency of their OTJs. The study used anonymous questionnaires and focus group interviews to gather data, which were collated and analysed by the researcher. Participants reported making OTJs using multiple sources of assessment evidence, most of which was age appropriate for students in their classes, with the exception of widespread running record use across all age groups. Challenges identified were: difficulties when making an OTJ with inconsistent evidence, working efficiently, gathering evidence of observations and conversations with students, and achieving consistent OTJs within and between classes. Participants were clear that more PLD was needed, especially to moderate and gain consistency of OTJs. These findings were consistent with recent ERO reports on National Standards implementation. The study was limited by its small sample size and geographic location. Further research is recommended across a wider sample to check the levelling of criteria in the Standards; and to explore examples of the evidence used in teachers’ reading OTJs and the process of relating that evidence to the Standards. In addition, clear guidelines and nationwide PLD are recommended to support teachers in gathering useful assessment evidence and making more reliable and consistent OTJs.
Acknowledgements

Thank you to my supervisor, Dr. Libby Limbrick for her recommendations and questions, clarity and precision, and wealth of knowledge in the teaching and assessment of reading.

Thank you to the participants of this study, whose willingness to share their practice in making reading OTJs made this research possible.

Thank you also to the many other educators, friends, and family members who have provided feedback, challenges, ideas and support.
# Table of contents

## 1. Introduction

1.1 Purpose of research 1  
1.2 Overview of thesis report 1  
1.3 Introduction of National Standards 1  

## 2. Review of the literature

2.1 Studies of National Standards implementation in New Zealand 6  
2.2 Assessment of reading in New Zealand 7  
2.3 Overall Teacher Judgement 11  
2.4 Moderation 14  
2.5 Teachers’ professional learning and development to assess with relation to the National Standards 15  

## 3. Method

3.1 Design 17  
3.2 Participants 18  
3.3 Data Collection 19  
3.3.1 Questionnaires 19  
3.3.2 Focus group interviews 19  
3.4 Data Analysis 19  
3.5 Validity 20  
3.6 Reliability 20  
3.7 Ethics 21  

## 4. Results

4.1 Overview 22  
4.2 Questionnaires 22  
4.2.1 Assessment tools and processes identified as used by participants in determining reading OTJs 22  
4.1.3 Teachers’ weighting of each assessment’s contribution to reading OTJs 26  
4.1.4 Sources of professional learning accessed to make reading OTJs 29  
4.1.5 Additional support identified for development in making reading OTJs 30  
4.1.6 Participants’ reported confidence in reading OTJs 31  
4.2 Focus groups 33
4.2.1 Use of information gained from assessment tools and processes 33
4.2.2 Advantages and challenges when using multiple sources of evidence 34
4.2.3 Weighting the importance of assessment tools and processes when making junior, middle and senior reading OTJs 35
4.2.4 Professional learning and development 36
4.2.5 Consistency and moderation of OTJs 37
4.2.6 New learning gained through making OTJs of students’ reading 38
4.3 Summary of results 39

5. Discussion
5.1 Assessments and processes used when making OTJs in reading 40
5.1.1 Use of conversations 40
5.1.2 Use of observations 41
5.1.3 Use of running records 41
5.1.4 Use of norm-referenced, standardised assessments 42
5.1.5 Challenges in using existing norm-referenced, standardised assessments 43
5.1.6 Triangulating evidence to make an OTJ 45
5.2 Professional learning and development 45
5.2.1 Use of professional learning and development resources in print and online 45
5.2.2 Use of professional learning and development workshops and courses 46
5.3 Confidence, consistency and moderation 47
5.4 Limitations 50
5.5 Recommendations 51
5.6 Conclusions 53

Appendices
1. Questionnaire 54
2. Focus group indicative questions 57
3. Assessment tools referred to in this study 58

Reference list 59
Chapter 1 Introduction

1.1 Purpose of research

The New Zealand Curriculum Reading and Writing Standards for Years 1-8 were published in 2009 (Ministry of Education); marking a significant shift in New Zealand’s educational policy and practice. Early studies of National Standards implementation reported differing results in teachers’ preparedness for and confidence in making reliable and valid overall teacher judgements (OTJs). These studies focussed predominantly on National Standards as a whole, rather than investigating OTJs in each area of the Standards. This research focuses on teachers’ use of assessment tools with relation to the reading National Standards, and examines issues particular to reading assessment within the New Zealand context.

In recent years there has been a strong emphasis on assessment and moderation in writing, and in the teaching of numeracy, but there has been little support for teachers to update their knowledge and skills in the teaching of reading. As a Resource Teacher: Literacy the researcher has experienced many enquiries from teachers asking for guidance on appropriate assessment tools and processes for students at different curriculum levels, and on making and moderating OTJs. This suggests that the reading National Standards pose particular difficulties for teachers because of insufficient recent professional learning in teaching and assessing reading. This research examines what information teachers of junior, middle and senior students use to make their OTJs for the reading National Standards.

1.2 Overview of thesis report

This report contains an introduction (Chapter one) and literature review examining studies of National Standards implementation and current literature on reading assessment in New Zealand (Chapter two). Chapter three describes the methodology of a research study designed to examine how teachers make their OTJs for the reading National Standards. The data collected in the study are reported in Chapter four, and discussed with relation to the literature in Chapter five. Implications of the data and recommendations are made at the end of Chapter five.

1.3 Introduction of National Standards

National standards of educational achievement are used in many countries with the common aim of measuring whether students are on track for achievement later in life. The results are put to a variety of additional purposes, including reporting to parents, determining whether schools are meeting targets,
and comparing school performance. The national standards of three English speaking jurisdictions were briefly examined for similarities and differences to New Zealand’s Standards. A key feature of other jurisdictions’ standards was that nearly all have national tests at stages of the curriculum with which to measure student achievement. England, Wales and Northern Ireland began introducing National Curriculum assessments in 1991, with the assessment schedule eventually testing students in years two, six, and nine. England and Northern Ireland’s assessments for year two and year nine students are now teacher assessments (for which schools may choose to include the use of standardised tools); and for year six both National Curriculum Tests and teacher assessments (Department for Education: Standards and Testing Agency, 2012, National Curriculum Assessment, n.d.). Wales now has teacher assessments for years two, six, and nine. Australia introduced NAPLAN (National Assessment Program for Literacy and Numeracy) in 2008. The programme uses national tests in literacy and numeracy to determine whether students are performing above, at or below the national minimum standards for their year level (National Assessment Program - Literacy and Numeracy, n.d.). The Common Core State Standards were released in the United States in 2010, with 45 out of 50 states adopting the initiative. The computer based assessments for these standards are in development by two organisations (Common Core State Standards Initiative, n.d.). The testing used in England, Northern Ireland, Australia, and in development in the United States, has been described as ‘high stakes assessment’ and has been criticised for disadvantaging students who do not perform well in tests, and narrowing the content of what is being taught or teaching to the test. An alternate view is that a national testing regime has the advantage of ensuring that all students can be assessed on the same content, and under similar conditions.

In New Zealand, national exemplars of reading and writing - which teachers could use to benchmark students’ progress against - were recommended a decade before the introduction of National Standards, in the 1999 Report of the Literacy Taskforce (Literacy Taskforce, p. 18). The taskforce also recommended that the “Government provide externally referenced assessment tools so that teachers and schools could monitor their progress towards the national goal of literacy for success by 2005” (ibid). The introduction of the Literacy Learning Progressions (Ministry of Education, 2010), and The New Zealand Curriculum Reading and Writing Standards for Years 1-8 (Ministry of Education, 2009) fulfilled the former recommendation of the Literacy Taskforce; and two new standardised, norm-referenced assessments of reading were developed in response to the latter recommendation. The Supplementary Tests of Achievement in Reading (Elley, Ferral, & Watson, 2011) were developed by the New Zealand Council for Educational Research (NZCER) to provide additional information about students’ reading competencies to that provided by the existing Progressive Achievement Tests: Reading (New Zealand Council for Educational Research, 2008); and the Assessment Tools for Teaching and Learning (asTTle) (Ministry of Education, n.d., c) were developed by the University of
Auckland under contract to the Ministry of Education to provide norm-referenced assessments that could be tailored to teacher and student needs. These tools gave New Zealand schools some options when selecting a standardised, norm-referenced tool to measure students’ reading progress over time.

The purpose of New Zealand’s National Standards is described in the introduction of *The New Zealand Curriculum National Standards for Reading and Writing for Years 1-8*:

> The National Standards provide a nationally consistent means for considering, explaining and responding to students’ progress in years 1-8. They provide reference points, or signposts, that describe the achievement, in reading, writing, and mathematics, that will enable students to meet the demands of the New Zealand Curriculum. They will help teachers to make judgements about their students’ progress so that the students and their teachers, parents, families and whanau can agree on the next teaching steps. (Ministry of Education, 2009, p. 4)

National Standards in New Zealand differ from those in the jurisdictions described above in that they provide achievement criteria for students, without providing specific assessment tools with which to measure them (in reading, the achievement levels are described through examples of appropriate text difficulty, accompanied by examples of appropriate comprehension and thinking skills relating to the text). Sewell (cited in Nicholson, 2010) describes the difference: “Unlike standards-based assessment in other countries, our standards do not rely on national testing. Instead there is an emphasis on teacher professional judgements, assessment for learning principles and practice, and the importance of information sharing to support student learning. This is a novel approach when compared with other jurisdictions” (pp. 3-4). Instead of using national tests to evaluate student performance, New Zealand teachers draw on many sources of informal and formal assessment information to make an OTJ of students’ achievement in relation to the National Standards. An additional challenge for educators and students is that New Zealand’s National Standards are unashamedly aspirational. Instead of meeting national minimum standards such as those in Australia, or achieving within an average band in norm-referenced tests such as in England and Northern Ireland, in New Zealand the “Standards are intended to raise achievement and promote the best possible progress for all students. Therefore, in some cases they establish an expectation that is higher than average” (Chapman, 2010). Assessing in relation to standards, making OTJs, and reporting to aspirational criteria are new processes for New Zealand teachers, and involve significant amounts of new learning and expertise.

Teachers in New Zealand primary schools have traditionally used an individualised, formative approach when reporting to parents on the progress students have made, and their next learning steps (e.g., Ministry of Education, 2006, p.71). National Standards have introduced a significant shift in that since February 2010, schools have been required to report to parents about their child’s achievement.
and progress using a four point scale of well below, below, at, or above the National Standard for their age or year level, in reading, writing, and mathematics.

Wylie and Hodgen (2010) describe this shift in educational policy:

The comparison of every child with a year-level (or years-at-school), and the need to place every child into one of the four categories of performance in relation to these benchmarks introduces a uniformity of expectation for every child that has not existed in the New Zealand education system for over 50 years, and which does not exist within the wider New Zealand curriculum, in which the National Standards fit (p.18).

With the introduction of National Standards, debate around reporting student achievement as well below, below, at, or above a National Standard has been both emotive and political. Concerns over labelling students, and the potential impact of league tables comparing schools’ achievement data have been prolific in the press. While academics in Colleges of Education and Universities have demonstrated public concern over the value of National Standards policy and implementation - with 170 academics signing an open letter opposing league tables based on National Standards - (Thrupp, 2012); some have defended the value of schools clearly stating students’ achievement with regard to benchmarks:

Research has shown that students can go through primary schooling thinking that they are positioned well for a prosperous secondary and tertiary career, only to find when faced with higher education curricula that they are significantly under-prepared” (Robinson and Timperley, 2004; cited in Hubbard & Annan, 2012, p.2).

Greaney (2009), in response to criticism that literacy and numeracy were singled out in the creation of National Standards, highlighted the importance of a clearer understanding of reading development and reading education in New Zealand. He writes: “It is unfortunate but true that the New Zealand ‘tail of underachievement’ is not in the arts, in technology, or social sciences. Rather it is in reading… This is where appropriate assessments could help schools identify such students and to help teachers develop appropriate programmes for them” (p.11).

Amidst political and media debate over the ethics of National Standards, there has been less emphasis on investigating the practical issues of whether teachers and schools have had sufficient knowledge and support to make valid and reliable assessments of student achievement with relation to these Standards. There have been no known studies specifically investigating the tools and processes used to make up an OTJ in reading. Anecdotal evidence suggests that this is a challenging area to evaluate and moderate as students’ reading does not necessarily produce tangible work. This contrasts with writing, in which a written piece can be evaluated by different teachers to get a shared understanding of the student’s
performance in relation to the National Standards. The lack of data on teachers’ preparedness to implement the reading National Standards is of concern. If teachers lack a clear understanding of appropriate assessment tools and processes, and practical skills in moderation, then their assessment will not be valid or reliable. This study attempts to gather initial data on teacher practice in implementing the reading National Standards, in order to identify areas for further professional development.
Chapter 2   Review of the literature

Studies of National Standards implementation in New Zealand were reviewed to gain an overview of Standards-based assessment. Large and small scale studies in New Zealand were reviewed to ascertain common themes and challenges in teachers’ assessment for the Standards. Specific research investigating the assessment of reading for the National Standards was sought, but little was available.

2.1   Studies of National Standards implementation in New Zealand

Three large organisations – New Zealand Maths Technology on behalf of the Ministry of Education; the New Zealand Council for Educational Research (NZCER); and the Education Review Office (ERO); conducted large scale studies to investigate school and teacher implementation of National Standards in 2010.

A contracted report by New Zealand Maths Technology to the Ministry of Education - *National Standards: School Sample Monitoring and Evaluation Project, 2010* (Thomas & Ward, 2011) investigated school and teacher readiness to implement the Standards by surveying the teachers, principals and Boards of Trustees from 104 schools. Thomas and Ward gathered data on making and moderating reading, writing and mathematics OTJs, the dependability of these OTJs and reporting to parents with relation to the National Standards. Teachers chose which area of the Standards to provide information about, and 66 teachers provided data on their reading OTJs.

NZCER used a similar survey-based approach in the *NZCER 2010 Primary and Intermediate Schools National Survey: A Snapshot of Overall Patterns and Findings related to National Standards* (Wylie & Hodgen, 2010). They gathered information about National Standards implementation from teachers, principals, and Boards of Trustees from 350 randomly selected schools, including data on making and moderating reading, writing and mathematics OTJs, reporting to parents, professional development, and the uses and impact of National Standards.

Thomas and Ward (op. cit.), and Wylie and Hodgen (op. cit.) found similar patterns in the sources of evidence used to make OTJs for the reading national standards, with Wylie and Hodgen presenting differences in assessment tools used by junior and senior teachers. The two studies reported different amounts of time used to make a reading OTJ, but found similar patterns of moderation practice with approximately half of the participants from each study reporting that they moderated reading OTJs in 2010.
The Education Review Office (ERO) published a series of reports over three years (Education Review Office, 2010, 2011, 2012a, 2012b) examining aspects of working with the National Standards. These reports contain more general information pertaining to the strengths and areas for improvement in National Standards implementation, rather than looking at OTJs in each area of the Standards. ERO initially investigated schools’ preparedness to work with the new documents. They identified the strengths of schools that were working well with the National Standards, and made recommendations to address the concerns of teachers and schools. The most recent of these reports (Education Review Office 2012a, 2012b) used data from 2011, and examined the extent to which schools were working effectively with the Standards.

A smaller scale study by Poskitt and Mitchell (2012) investigated teachers’ understandings about making OTJs in 2010. These researchers distinguish between an analytic approach to OTJs – in which data from assessments are weighted, scored and totalled to form an OTJ – and a holistic approach – for which evidence is gathered and considered as a whole against the Standard. The quantitative studies described earlier use aspects of the analytic approach to understand how large groups of teachers make OTJs – for example, by setting particular criteria for quality OTJs (Education Review Office, 2012a), and by weighting the importance of assessment tools (Thomas & Ward, 2011). Poskitt and Mitchell theorise that: “Bringing together multiple samples of work in order to derive a holistic judgement in relation to the National Standards is the essence of the New Zealand OTJs” (p.58). In their qualitative research they found that teachers had varied understandings of OTJs, of useful supporting evidence, and of moderation processes. They describe making a holistic judgement of achievement with relation to a Standard as a challenging task, but one that may derive rich formative information and develop deep teacher learning over time. These researchers emphasise the importance of moderation in this process.

The studies outlined above together with a number of smaller studies and reports, are discussed in more detail with relation to specific topics later in this dissertation.

2.2 Assessment of reading in New Zealand

Reading assessment in New Zealand schools has followed a formative approach, with the focus for assessment data strongly on the next steps for teaching and learning. The Ready to Read and School Journal resources (published by Learning Media for New Zealand schools) are levelled, with benchmark levels for different age groups, however Effective Literacy Practice in Years 1-4 (Ministry of Education, 2003) (the current teacher handbook for junior literacy) uses general terms to describe acceptable progress:
“There is a widely held expectation among teachers that children at the end of year 1 will be reading at (or beyond) Blue to Green levels on the Ready to Read colour wheel (p.71)”. And: “There is an expectation among teachers that children at the end of year 4 will be able to read, comprehend, and respond to texts that are widely agreed to be appropriate at their chronological age” (p.73).

This emphasis on teachers’ existing knowledge implies that the expectations for junior reading development may have been tacit rather than explicitly defined in print. A more recent publication - *The Literacy Learning Progressions* (Ministry of Education, 2010) is more direct in clarifying the expectations for the expected reading text level, and decoding and comprehension skills expected for students at each year level of school. In many ways, this document fulfils the recommendation of the Literacy Taskforce for national exemplars of reading and writing.

The Literacy Taskforce recommended that the government provide externally referenced assessment tools in literacy and, while there is no universal tool that all schools use; most schools choose to use one or more of the following norm-referenced and standardised tools (please see appendices for a brief description of the assessment tools referred to in this study). *The Observation Survey* (commonly known as the six year net) (Clay, 2002) is a series of tasks designed to screen for the lowest achieving six year olds in order that they may qualify for a place on Reading Recovery. The *Progressive Achievement Test: Reading Comprehension and Vocabulary* (PAT) (New Zealand Council for Educational Research, 2008) was originally developed in the 1960s to check on students’ progress in relation to the curriculum. The PAT assesses students’ independent reading skills. The *Supplementary Test of Achievement in Reading* (STAR) (Elley, Ferral, & Watson, 2011) was developed between 1999 and 2003 to provide additional information about students’ independent reading skills. Both PAT and STAR have undergone recent revision. Parr, Reddish and Timperley (2007) identified ceiling effects when using the previous edition of STAR with senior students and it is not known whether the new edition has overcome this. *asTTle* (Assessment Tools for Teaching and Learning) (Ministry of Education, n.d., c) are a series of norm referenced assessments developed by the University of Auckland under contract to the Ministry of Education. Access to asTTle is free to all schools and the online assessments enable teachers to design tests to measure specific skills.

The formal tools described above are a small part of the number of formal and informal assessment tools available to New Zealand teachers. *The New Zealand Curriculum Reading and Writing Standards for Years 1-8* (Ministry of Education, 2009) recommend that teachers use the assessment
tools and processes (including those described above) listed in the two volumes of *Effective Literacy Practice* (Ministry of Education, 2003, 2006), and draw the information together to make an OTJ. “Making an overall teacher judgement will mean drawing on evidence gathered up to a particular point in time and analysing it in order to make an informed, balanced judgement in terms of the student’s actual performance - how it lines up with what is expected in terms of the relevant standard” (Ministry of Education, 2009, p. 13).

The Ministry of Education’s *Position Paper on Assessment [Schooling Sector]* (2011) describes the need for all involved in education to step up their assessment practices to ensure that they are effective. Timperley and Parr (2010) discuss how working with the National Standards has the potential, both to sharpen teachers’ assessment practices, and to promote ongoing learning and sharing in the teaching community about moderation. They caution that the assessments used when making an OTJ against the National Standards “must be fit-for-purpose and allow appropriate diagnoses to be made” (p.13). A fit for purpose assessment is one that gathers useful information about a student’s abilities, and indicates areas that require instruction. Teachers need a deep pedagogical knowledge of reading development to understand which assessment tools and processes should be used at particular stages, and which will be most appropriate for summing up students’ achievement.

*The New Zealand Curriculum Reading and Writing Standards for Years 1-8* (Ministry of Education, 2009) distinguishes between the early reading years (years one to three) when students are building the foundations of literacy through guided reading of texts on the *Ready to Read* colour wheel, and the later years (years four to eight) when students are reading an increasing variety of texts across the curriculum (pp. 9-10). Differentiated assessment tools and processes are indicated for these stages of reading. The Standards for years one to three include a seen text accuracy level of 90% at the appropriate level on the colour wheel as a criterion, and cognitive behaviours that may be developed in subsequent readings. The Standards for years four to eight give examples of appropriate cognitive reading behaviours for students at each level, and stress that these must be observed on a variety of texts at the appropriate level. Students initially learn to read through teacher-scaffolded instruction and, over time, develop the independent reading skills and strategies needed for secondary education. Appropriate assessment for older students should include some measures of independent work to capture evidence of this development. In order to capture evidence of learning at different stages of reading, one would expect to see a difference in the fit-for-task tools used to assess reading in junior, middle and senior classes.

Appropriate reading assessment tools for students after one year at school include the norm-referenced Observational Survey (Clay, 2002) and Burt word test (Gilmore, Croft, & Reid, 1981), observations of
students’ reading behaviours, and a running record on a seen text (that is, the student has encountered the text during guided reading, and the running record is taken in the next few days). Running records are a common procedure for assessing young students’ oral reading accuracy and behaviours. Nicholson (2010) investigated the validity of using running records as an assessment measure by comparing unseen running record results with norm referenced measures of word reading, phonological awareness, decoding, and invented spelling on 101 year two students. Nicholson found that the measures of word reading, decoding and invented spelling correlated highly with students’ performance in running records, and that the measure of phonological awareness correlated significantly with students’ performance on running records. Nicholson’s study suggests that running records are a valid and reliable assessment for young students who are learning to read.

Timperley (2004) found that schools were routinely administering running records across all year groups, often in order to fulfil school reporting requirements. A small scale study conducted across a cluster of rural schools in 2011 (Quick, 2011) found the practice of using running records as an assessment for all students in the senior primary years was widespread, despite Effective Literacy Practice in Years 5-8 clearly stating that: “Running records are not intended for use with fluent and independent readers, whose development needs probably relate to making deeper meaning and thinking more critically about texts” (Ministry of Education, 2006, p. 60). More appropriate assessments for fluent readers include standardised assessments, such as PAT, aSSTle and STAR to provide information about students’ independent reading comprehension, Assessment Resource Bank (ARBs) tasks to provide information about students’ comprehension in group and independent settings, and observation of students’ cognitive behaviours and conversations as they read across the curriculum. When investigating early National Standards implementation, Wylie and Hodgen (2010) noted a shift in the kinds of reading assessments used between junior and senior classes, with middle and senior teachers more frequently using standardised assessments (p.33). They also found that nearly all teachers ranked running records as being of high importance when making reading OTJs. While ‘fit for purpose’ assessments are essential in making reliable and consistent OTJs for the reading National Standards, the degree to which teachers are using age appropriate tools for the level they teach is not clear.

2.3 Overall Teacher Judgement

The process of making an overall teacher judgement to determine student achievement against the reading Standards should involve a process of triangulation; using learning conversations, observations of students as they engage in reading activities, and outcomes from assessment tools (Ministry of
The assessments outlined above are grouped under the three main categories in the diagram below to show the triangulation process used when making an OTJ.

**Figure 1.** Overall teacher judgement

(Ministry of Education, n.d., b)

Learning conversations (referred to in this study as ‘student voice’ in order to capture verbal and non-verbal dialogue with students) are a key characteristic of assessment within *The New Zealand Curriculum*, (Ministry of Education, 2007), which states that students are expected to “discuss, clarify, and reflect on their goals, strategies, and progress with their teachers, their parents, and one another.” (p. 40). The inclusion of observation of process as a key area in gathering evidence for the National Standards reiterates the importance of on-going formative practice and assessment for learning; in which teachers evaluate students’ strengths and weaknesses as they learn, and plan for teaching experiences to support further learning. Using evidence from tool outcomes provides a check on both students’ performance with others of the same age/year level, and on the reasonableness of teachers’ formative judgement. *Te Kete Ipurangi* (TKI), the online resource for New Zealand teachers recommends that targeted conversations and observations should occur most frequently in classes, tasks should occur moderately frequently, and use of standardised tools should occur less frequently. Each of the three areas should be supported by the teacher’s tacit knowledge of the learner, and by exemplars and illustrations of the standards (Ministry of Education, n.d., d). An OTJ is not simply a summary of these sources of evidence, but a careful consideration of the evidence elicited through the use of appropriate assessment tools in relation to the National Standards.
While teachers are asked to continue with current good assessment practices when making an OTJ, the process of pulling data from various sources together in order to make a Standards-referenced judgement of student achievement is a new one for teachers. Hubbard and Annan state: “Teachers had routinely used both objective and subjective data to assess individual student achievement, but as one principal explained, *OTJ was not a term or practice in common use before the advent of national standards*” (2012, p. 4). Teachers are now required to review their evidence to make a summative OTJ with relation to a Standard, but the number of sources needed to make a reliable and efficient OTJ has been left up to the interpretation of schools. Wylie and Hodgen (2010) found that most teachers making OTJs used four or more kinds of assessment evidence. Some teachers were using many more, with 42% of teachers using six or seven sources of evidence, and 18% of teachers using eight or nine sources. Wylie and Hodgen state “It would be useful for more qualitative study of the value of having so many sources feeding into an OTJ” (p.32). Considering evidence from a variety of sources takes time, and making an OTJ needs to be efficient in terms of teacher workload, and well as providing a robust assessment of student achievement. Thomas and Ward (2011) found that approximately half of the teachers surveyed took up to ten minutes to make an OTJ, which they considered efficient (p. 1). An alternate and longer estimate comes from Wylie and Hodgen (op cit.), who found that teachers took approximately 20 minutes to make a reading OTJ (p.34).

ERO investigated schools’ use of assessment data in making OTJs (for all National Standards rather than specifically the reading Standards) between 2010 and 2011 and found that schools were at different stages of using assessment data effectively to make OTJs. Their first review, conducted in the first half of 2010 (Education Review Office, 2010) found that while 88% of teachers were using assessment data to inform their teaching, “the use of data was highly variable” (p.9). For the 20% of schools that were not well prepared to work with the National Standards a common challenge was the “need to improve current assessment practice through the use of a range of assessment tools in reading, writing and mathematics” p.20). Eighteen months later ERO (2012a) reported that the 19% of schools that were not working well with the Standards continued to experience difficulties with assessment including “developing processes to gather relevant and useful achievement data [and] increasing teachers’ confidence in using multiple sources of evidence to make OTJs” (p.8).

Thomas and Ward (2011) and Wylie and Hodgen (2010) investigated National Standards in the first year of implementation. Both studies found that teachers were using a variety of sources to make an OTJ in reading, and most of those sources were considered relevant to assessing reading for the National Standards. Thomas and Ward reported that teachers regarded classroom observations as the most important source of evidence when making an OTJ, but noted that: “The observations described by teachers tended to be general in nature rather than describing students’ particular abilities in a way
that might be considered informative in terms of OTJs” (p. 1). As observations of students’ reading contribute a large proportion to OTJs, this finding is of concern.

To add to the complexity of making an OTJ in reading, teachers are required to assess student achievement as students read across the curriculum, rather than simply in instructional reading sessions. Teachers are reminded to “specifically consider how well each student is using reading and writing as interactive tools to enable them to learn in all curriculum areas” (Ministry of Education, 2009, p. 5). As ERO found that nearly one fifth of schools were having difficulty in using a range of assessment data effectively, the above requirement may be too complex and sophisticated for schools developing appropriate assessment processes. Weal and Hinchco (2010) regard assessment for National Standards using data gathered across the curriculum as part of a three stage developmental continuum for teachers; with assessment in the specific domain (in this case, during instructional reading) being the initial stage, and dynamic assessment in authentic integrated curriculum experiences being the third and most complex stage. It is likely - given that only 22% of schools were found to be working well with the Standards in 2011 (Education Review Office, 2012a) - that many schools and teachers will be at the first stage in this continuum, that is, assessing reading primarily during reading specific tasks.

While researchers (e.g. Cameron, 2010; Hubbard & Annan, 2012; Poskitt & Mitchell, 2012) identified many challenges for teachers when making OTJs, participants from two early studies reported high levels of confidence, and moderate levels of consistency in their OTJs. Thomas and Ward (2011) found that teachers and principals were very confident in the accuracy of their OTJs across all National Standards, and in the consistency of OTJs across their schools (p.1). Wylie and Hodgen (2010) found that 38% of teachers making OTJs thought there was a high level of consistency in OTJs across the school (p.36). They also noted an interesting shift between the confidence in reading OTJs from teachers of New Entrant students to teachers of senior students, and suggested that:

The difference could be because early reading is about the acquisition of skills rather than applying skills to a widening range of material, and because there has been a long tradition now in New Zealand of graded readers for the early school years, and assessments associated with the 6-Year net (p.32).

Self-reported confidence in one’s OTJs does not appear to translate to reliable OTJs, as ERO found that only 25% of teachers met the criteria for working well with the Standards in 2011 (Education Review Office, 2012a, p. 9). It is possible that ERO’s use of specific criteria for good practice in making OTJs has shown the task to be more challenging than teachers initially believed.
2.4 Moderation

Consistency is a challenge when using OTJs to assess in relation to Standards. “Moderation is the process of teachers sharing their expectations and understanding of Standards with each other in order to improve the consistency of their decisions about student learning” (Ministry of Education, n.d., a). Secondary teachers are routinely involved with moderation to ensure consistent assessment of student achievement for national qualifications. Moderation is a newer concept for primary teachers, particularly in the area of reading. When reporting on moderation practices and challenges, ERO (Education Review Office, 2011) found that confident schools had accessed professional learning and development in moderation, had a reflective and open school culture that allowed for professional discussions, and looked on moderation as an area for on-going development.

In 2010, schools were beginning to moderate OTJs but Wylie and Hodgen (2010) found that teachers “were more likely to have experienced this [moderation] in relation to school benchmarks or progressions in writing (77 %), than in relation to reading (55 %) or mathematics (53%) (p.35). When investigating moderation, Thomas and Ward (2011) found that approximately half the teachers surveyed participated in within-school moderation of reading OTJs. Fewer schools were engaged in any between-schools moderation, and most of this moderation was focused on writing (p.1).

Poskitt and Mitchell (2012) argue that “professional conversations within moderation processes are the basis of potential deep learning and shared understandings of curriculum, pedagogy and assessment” (p.72). They acknowledge that moderation is a process that takes time to develop but believe that it is possible to achieve greater consistency over time. Wylie and Hodgen (2010), and Thomas and Ward (2011) found that teachers who had engaged in moderation had a higher level of confidence in the consistency of OTJs across the school to that of teachers in schools that did not moderate their OTJs, suggesting that the shared task of moderation helps teachers to trust the OTJs of their colleagues.

It appears that improving the consistency of OTJs and developing reading moderation processes are key areas in which schools and teachers need support. The most recent study by ERO (Education Review Office, 2012a) found that:

The most commonly cited challenge with implementing the National Standards was concern around the accuracy and moderation of overall teacher judgements. This included teachers’ confidence in making judgements, the quality of the achievement information used to make them, and the alignment of various assessment tools with the standards. In some schools, the teachers’ capability to make judgements and carry out moderation varied widely across the school (p.17).
2.5 Teachers’ professional learning and development to assess with relation to the National Standards

There has been no nationwide National Standards professional learning and development (PLD) programme offered for teachers. While schools have undertaken significant professional development in the assessment and moderation of writing in recent years, the assessment of reading has had less coverage. Consequently the assessment of reading across New Zealand is likely to be variable. The current means of access to PLD in New Zealand schools is either by self-directed or school-directed study and research of resources on National Standards, or through literacy advisors working for Universities, Colleges of Education, and independent companies. Schools choose to access the support of these advisors, which has meant that all teachers do not necessarily have the same opportunities for access to PLD. Boards of Trustees and principals who are ethically opposed to National Standards are unlikely to request advisory support in implementing them. Small and remote schools are less likely to be able to access PLD than larger schools in cities with Universities and other companies offering courses and on-going development.

Print resources to support teachers when making reading OTJs include The New Zealand Curriculum: Reading and Writing Standards for Years 1-8 (Ministry of Education, 2009), Effective Literacy Practice in Years 1-4, and Years 5-8 (Ministry of Education, 2003, 2006), The Literacy Learning Progressions (Ministry of Education, 2010a), resources online at TKI (Ministry of Education, 2010b) and alignment documents to support teachers’ interpretation of the Observation Survey (6 year net), STAR, PAT, and asTTle results with relation to the reading standards (Ministry of Education, 2010c). Other publications offer examples to support teachers in making OTJ, e.g. an Education Gazette (Ministry of Education, 2010 d) article demonstrates the triangulation process, and gives examples of informal and formal assessment tools for years one, four and eight. The availability of these resources, particularly those accessed online, may not be widely known by teachers. It is possible that accessing material via the internet is not ingrained in many schools’ cultures, and that schools are not aware of updated documents, recommendations and exemplars.

When teachers accessed PLD through courses, workshops or advisors; they reported mixed experiences (e.g., Cameron, 2010, Education Review Office, 2010, Wylie & Hodgen, 2011). The 2010 ERO report found that most schools that had accessed outside support to understand the standards were positive about the impact this advice had. This included external advisors as well as Ministry of Education support. However the report cautions that PLD in this area needs to be ongoing (p.17). Other studies found that many teachers experienced the training they had accessed to be inadequate or not helpful, e.g., Wylie and Hodgen (2010) found that a large proportion of teachers who had accessed professional
development either funded by the Ministry of Education (69%) or by an outside provider (60%) said that it had little or no impact on their practice. ERO’s more recent report (Education Review Office, 2012a) noted that teachers in a few schools found the PLD they had accessed to be confusing or unhelpful. Cameron (2010), in her small scale investigation of National Standards implementation also reported concerns from teachers that the professional information and learning provided to schools to implement the National Standards was not adequate: “The Ministry of Education professional development in relation to the Standards was not seen as consistent or helpful in terms of communicating the intent of the Standards or helping teachers strengthen their judgements of student achievement” (p28).

While early findings as to the efficacy of PLD opportunities were mixed, more recent case studies of schools engaged in good practice with the National Standards (Education Review Office, 2012b) consistently cite accessing PLD as pivotal in working well with the Standards: “PLD increased their understanding and awareness of the standards and of student achievement across the school and within their class” (p. 11). Conversely, features of schools that were not working with all the requirements of the National Standards were: difficulty with access to relevant and useful PLD; guidance as to which assessment tools to use, and knowledge and understanding of making and moderating OTJs. Without universal access to quality PLD, it appears that small, remote and struggling schools may continue to experience difficulties in working with the National Standards; and may not have the means to improve their practice.

Schools have had the opportunity to work with and report in relation to the reading and writing Standards for more than two years. While large scale studies of early implementation indicated generally appropriate use of assessment tools and confidence in OTJs, the qualitative data from smaller studies suggests that teachers were challenged by assessment in relation to the Standards. All known studies investigated Standards implementation in general, with little data provided on teachers’ competencies and challenges in each of the reading, writing, and mathematics Standards. ERO’s ongoing studies reveal a pattern in which the proportions of schools working well with the Standards and having difficulty with the Standards are not altering significantly over time. One wonders whether the teacher support materials and PLD opportunities available are adequate for teachers to make robust judgements of student achievement with relation to the Standards, or whether more effective guidance and training is required on assessment for the National Standards, and in making and moderating OTJs. This study attempts to answer these questions by investigating how teachers make their overall teacher judgements for the reading National Standards.
Chapter 3  Method

3.1  Design

The study used a mixed methods design (Cresswell, 2005) in order to gather quantitative data about teachers’ assessment practices, and to gather qualitative data from a representative teacher (literacy leader) from each participating school about how and why they used particular assessment tools and processes to make reading OTJs. This design was selected to provide statistical data on trends in teachers’ implementation of the reading National Standards, and to provide qualitative interview data to describe teacher practice and provide a multifaceted perspective. The questionnaire enabled a wider coverage of respondents which was complemented by a greater depth of response within the focus groups. A mixed methods approach is particularly relevant to this study as its aim is to investigate the process and challenges of gathering data from different sources. The study uses an analytic approach to investigate teachers’ OTJs for the reading National Standard, similar to that used by Thomas and Ward (2011) and Wylie and Hodgen (2010). This is to enable quantitative data to be gathered about different reading assessments and their contribution to OTJs.

The research questions were:

- What assessment tools and processes do teachers use to make OTJs of reading achievement with relation to the National Standards?
- What proportion does each assessment tool and process contribute to teachers’ OTJs?
- What professional development opportunities have teachers accessed to support their work with the reading National Standards, and what kinds of professional support do teachers require?
- Does the curriculum year level taught impact on the kinds of tools and processes used to assess reading?
- How confident are teachers in making OTJs for the reading National Standards?

A questionnaire was designed to gather information about the assessment tools and processes teachers used to make their OTJs for the reading National Standards. It included questions about the participants’ year level taught, their years of New Zealand teaching experience, the assessment tools and processes used when making a reading OTJ, the PLD accessed, and preferred for future learning, and confidence in OTJs (Please see appendices for questionnaire).

The questionnaire was trialled amongst teachers who were not involved in the study. Participants in the trial provided feedback as to whether the questions in the questionnaire accurately measured how teachers assess students’ reading in relation to the National Standards. In response to feedback from the
trial group, alterations and additions were made to ensure that the questions accurately measured how teachers assess students’ reading with relation to the National Standards.

The tools and processes listed in the questionnaire were those recommended in *Effective Literacy Practice in Years 1-4* and *Effective Literacy Practice in Years 5-8*, (Ministry of Education, 2003, 2006) and in the teacher support materials on TKI (Ministry of Education, 2010b). Other tools and processes were added in by recommendation of the trial group. Tools listed included those used for formative assessment, and for making summative assessments of different levels of students. Lists of items in the questionnaire were presented alphabetically to avoid bias. (Please see appendices for a brief description of the assessment tools referred to in the study)

Indicative focus group questions were formulated to gather more descriptive data than was possible in a questionnaire. These questions were designed to facilitate discussion about: assessment tools and processes, making an OTJ, the advantages and disadvantages in using multiple sources of evidence, the sources used for professional learning, moderating OTJs, and the next steps for teachers and schools with reading OTJs (Please see appendices for indicative focus group questions).

3.2 Participants

The sample of primary teachers - teaching years one to eight, and with a range of teaching experience - were a convenience sample of the school cluster (geographical and Ministry of Education funded) in which the researcher works. The 17 schools in the cluster ranged from one teacher schools in isolated areas, to larger town schools of up to 10 teachers. Information about the project, and consent forms to participate, were sent to schools in the last week of May 2012.

The Principal and Board of Trustees of each school were approached in writing and asked for consent to invite teachers in their school to participate in the study. Of the 17 schools in the cluster, 14 Principals consented to their teachers participating, should they wish to do so.

Approximately 59 teachers at the 14 participating schools were sent information about the study and were invited to participate by filling in the questionnaire and returning it by post.

In addition to participation via questionnaire, the literacy leader from each school was invited to participate in a focus group interview, probing the reasons for their reading assessment with relation to the National Standards.
3.3 Data collection

3.3.1 Questionnaires

Thirty five questionnaires were received from participants teaching year levels one to eight in one of the 14 participating schools. Participants completed the anonymous questionnaires in their own time during June 2012.

3.3.2 Focus group interviews

Six literacy leaders (teachers with particular responsibility for literacy in their schools) participated in the focus group interviews, which were held as three discussions with between one and three teachers participating in each discussion. It is likely that these literacy leaders had also contributed quantitative data via questionnaire, but as the questionnaires were returned anonymously, this cannot be verified.

Before the interview participants were provided with the indicative questions for discussion and a copy of the assessment tools and processes listed in the questionnaire. The focus group interviews took place in the last week of July and the first week of August 2012 and were held after school at the researcher’s office.

3.4 Data analysis

Questionnaire data was grouped into junior, middle and senior levels to reflect the multilevel classes most teachers taught. Year three data was considered junior for a year two-three composite class, or middle for a year three-four composite class. One participant taught across junior, middle and senior levels and their data was assigned to a best fit level of middle, as reporting it separately may have inadvertently identified the participant to the rest of the cluster.

The questionnaire data was analysed using descriptive statistics. The data on assessment tools and processes were clustered under the three types of assessment evidence (conversations with students, observations of students, and assessment tools) shown in figure one. The tools and processes used by teachers of different ages were compared. The transcribed focus group interviews were analysed for patterns and key ideas, and categorised under themes (Cresswell, 2009). Because of the small nature of the study, time restraints, and financial restraints all data was entered, coded and grouped by the researcher.
3.5 Validity

The questionnaire was trialled amongst experienced teachers to ensure that the questions were a valid measure of teachers’ knowledge, skills and practices when implementing the reading National Standards. Feedback from the trial group suggested that the questionnaire did measure these and therefore had internal validity.

The focus group questions were open ended to enable authentic discussions about reading OTJs. Literacy leaders were invited to add additional information about their work with the reading National Standards to ensure that the data gathered was an accurate description of their OTJs. “Validity… is one of the strengths of qualitative research” (Cresswell, 2009, p. 191). The focus group interview data has strong internal validity, that is, it accurately portrays the views of those interviewed. Conversely, this qualitative data has limited external validity as it is particular to the participants.

The size and geographical area of the sample mean that the study has limited external validity. Larger scale sampling would be necessary to determine whether the findings are typical of New Zealand teachers.

3.6 Reliability

The reliability of data gathering tools (in this case the questionnaires and focus group questions) is usually determined by whether their findings can be replicated by other researchers and on multiple occasions. As making OTJs in reading is a relatively new process for all teachers, one may find changes in the data gathered if this study was replicated. This study investigates how participants made their reading OTJs in the middle of 2012.

The questionnaire data was entered and checked to ensure no entry errors had been made. The interview transcriptions were checked and sent to the participants for checking, and, in two cases, to clarify wording that was unclear from the recording. The reliability of the data analysis may be compromised by the data entry and analysis being carried out by the researcher. The use of another researcher to gain inter-rater reliability was not feasible in a study of this size.
3.7 Ethics

Participation in the study was informed and voluntary. Boards of Trustees, principals, and teachers from participating schools were informed in writing about the purpose of the research and what involvement in the study entailed.

Principals were asked to distribute the questionnaires to possible participants in order to preserve anonymity. The questionnaires were anonymous and were returned via post to the researcher. The questionnaire was short and took approximately 15 minutes to complete. Participants were assured that no region, school, or teacher would be identified in any publication or workshop arising from the research.

Literacy leaders gave written consent to participating in a recorded focus group interview. They were asked to respect the confidentiality of other focus group participants, and were assured that their contributions would be reported anonymously. They were informed at the beginning of the focus group that they may ask for the voice recorder to be turned off at any time, and that they may leave at any time. The researcher (who chaired the interviews) planned to redirect the discussion should contentious debate arise, however this was not necessary. The focus group interviews took between thirty to sixty minutes.

As Resource Teacher: Literacy in the cluster, the researcher had a professional relationship with the participants. Therefore the participant information sheet clearly stated that schools and teachers would not be disadvantaged by not participating in the research, to ensure that no conflict of interest could be perceived.

Ethical approval was sought from the University of Auckland Ethics department and granted in May 2012.
Chapter 4 Results

4.1 Overview

The results are reported with the quantitative, questionnaire data followed by the qualitative, focus group data. The topics investigated were the assessment tools and processes teachers of different levels reported using, and the proportion these tools and processes contributed to OTJs, sources of professional learning utilised, preferred sources of future professional learning, and confidence in OTJs. Trends arising from the data were: that triangulating data from multiple sources was an established practice, and confidence when making reading OTJs was moderately high. Weighting the importance of different assessment tools and processes was more challenging. Participants identified a strong need for more PLD, and consistency and moderation were the major challenges identified in the focus group interviews.

4.2 Questionnaires

The questionnaire asked participants to select the school year level they taught: 14 teachers taught years one to three (junior), 13 teachers taught years three to five (middle), and eight teachers taught years six to eight (senior).

Participants were asked to select the number of years teaching experience in New Zealand schools they had had. Seven teachers had zero to five years teaching experience (novice), 15 teachers had six to 20 years teaching experience (experienced), and 13 teachers had over 20 years teaching experience (extremely experienced).

4.2.1 Assessment tools and processes identified as used by participants in determining reading OTJs

The questionnaire listed 18 possible sources of assessment information. Participants were asked to select the sources they used when making an OTJ of student achievement in relation to the reading National Standards. These are first listed to show the number of sources teachers used.
Table 1

*Mean Number of Sources of Evidence Used to Make a Reading OTJ (n=35)*

<table>
<thead>
<tr>
<th></th>
<th>Mean no. sources of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior</td>
<td>8.36</td>
</tr>
<tr>
<td>Middle</td>
<td>8.77</td>
</tr>
<tr>
<td>Senior</td>
<td>9.25</td>
</tr>
<tr>
<td>Novice</td>
<td>9</td>
</tr>
<tr>
<td>Experienced</td>
<td>8.8</td>
</tr>
<tr>
<td>Extremely Experienced</td>
<td>8.46</td>
</tr>
</tbody>
</table>

Teachers reported using between five and 14 sources when making an OTJ. The mean number of sources used by all participants was 8.71, and there was little difference between the means for level taught or experience, although those working in senior classes and novice teachers appeared to use a slightly wider range of assessment tools.

The sources of assessment information listed were grouped under the three categories recommended in TKI’s OTJ model (see figure one): student voice, classroom observations and informal assessment tasks, and standardised, norm-referenced assessments. The student voice category included conversations with students about their reading and information about their recreational reading, and the standardised, norm-referenced assessment category included the Observation Survey (six year net), the Burt word test, and the STAR, PAT and asTTle tests of reading. The observations category contained 11 options, including different types of running records, classroom tasks and informal assessments, and observations during guided reading and across the curriculum.
All participants reported using observation during guided reading time when making an OTJ. Nearly all used conversations with students (97% of all participants), and observation of students reading across the curriculum (83% of all participants) to inform their OTJs. Observations of students reading across the curriculum and information on recreational reading were higher for teachers of senior students (100%) than for teachers of junior (observations across curriculum=71 %, information about recreational reading=50%) or middle students (reading across the curriculum=85%, information about recreational reading=77%). Teachers of senior students were most likely to use their own (teacher made assessment) (63%), but least likely to use ARBs (13%) which was surprising.

Other sources listed by teachers included Magic tree books (a programme to learn letters and sight words, no reference available), the PAT listening test, graphic organisers from Teaching Reading Comprehension Strategies: A Practical Classroom Guide (Cameron S., 2009), and school learning indicators. Thirty-six percent of junior teachers and 8% of middle teachers reported using other sources. Some of these tools appear to be of formative, diagnostic use to the teacher, though the graphic organisers can display evidence of cognitive reading behaviours relevant to the Standards. As
the questionnaire was anonymous it was not possible to probe participants for their reasons for including particular tools.

Running records were frequently used across all levels. The questionnaire listed six types of running records which can be broadly divided into seen text (that is, the student has encountered the text during guided reading, and the running record is taken in the next few days); and unseen text (for which the student has not received instruction on the text). Various informal prose inventories such as PM Benchmark (Nelson Primary, n.d.), PROBE (Parkin & Parkin, 2011) and PROSE (Ayrey, n.d.) use unseen text running records accompanied by questions of comprehension asked by the teacher and answered orally by the student.

![Bar chart showing percentage of teachers using running records when making OTJs (n=35).](image)

Figure 3. Percentage of teachers using running records when making OTJs (n=35).

All participants reported using at least one kind of running record to make their reading OTJs and many participants used two or more kinds of running records. A large proportion of participants used seen text running records (91%). The use of seen text running records was higher for teachers of junior students (100%) than for middle (92%) or senior (75%) students but was still surprisingly high in the middle and senior school. In addition, all senior teachers reported using running records from informal prose inventories - PROBE, PROSE or PM. One teacher annotated the questionnaire with “Used when and if student shows areas of concern” to explain their use of PROSE and PROBE with senior students.
Of the 35 participants, 33 used some form of norm-referenced, standardised assessment when making their reading OTJs. The Observation Survey (commonly known as the six year net) (64%) and Burt reading test (71%) were frequently used by junior teachers. The Burt was used by nearly half of middle and senior teachers, and the six year net continued to be used by one middle and one senior teacher.

The most commonly used standardised test was PAT (60% of all participants). All senior teachers reported using this tool. The use of standardised tests of comprehension increased with year group level, with few junior teachers using PAT (14%) or STAR (36%), and more middle teachers using PAT (79%), STAR (57%), and asTTle (15%). All senior teachers reporting using PAT, and half reported using STAR. Few teachers (6% of all participants) used asTTle, which was surprising.

To summarise, participants used a range of assessment evidence from three broad areas to gather evidence for their OTJs.

4.1.3 Participants’ weighting of each assessment’s contribution to reading OTJs

Participants were asked to weight the assessment tools and processes they selected, by estimating the percentage each contributed to their overall teacher judgement of student reading achievement.

Thirty-two teachers answered this question and 23 (eight junior, eight middle, and seven senior) of these provided responses in the questionnaires in which the estimated percentages totalled 100%. These data are reported below (the other nine questionnaires had estimated percentages with a range of 40%-300%).

Figure 4. Percentage of teachers using norm-referenced, standardised assessments when making OTJs (n=35).

To summarise, participants used a range of assessment evidence from three broad areas to gather evidence for their OTJs.
The data were clustered under the three types of information teachers triangulate when making an OTJ: student voice, observations (with a subsection of running records), and norm-referenced, standardised tests. The mean percentage of each data cluster was calculated for junior, middle and senior teachers.

Table 2

*Range and Mean of Percentage Apportioned to Types of Assessment (n=23)*

<table>
<thead>
<tr>
<th></th>
<th>Student voice</th>
<th>Observations (Running records)</th>
<th>Norm-referenced tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior (n=7)</td>
<td>Range</td>
<td>0-20%</td>
<td>65-95%</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>9.375%</td>
<td>76.875%</td>
</tr>
<tr>
<td>Middle (n=8)</td>
<td>Range</td>
<td>5-30%</td>
<td>40-80%</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>13.125%</td>
<td>63.375%</td>
</tr>
<tr>
<td>Senior (n=8)</td>
<td>Range</td>
<td>10-35%</td>
<td>50-70%</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>18.57143%</td>
<td>59.71428%</td>
</tr>
</tbody>
</table>

There were large differences in the proportions allocated to each cluster type. Observations (including classroom activities and informal assessment tasks) accounted for a very large proportion of teachers’ OTJs (between 40%-95%).
Figures 5-7. Mean percentage each tool type contributed to OTJs of junior, middle and senior students’ reading (n=23). (Percentages rounded to nearest whole number)
Most teachers weighted several tools or processes with low proportions, indicating that they were not depending on the results of one tool. However once the tools were clustered by type (according to the TKI diagram, figure one), clear patterns emerged about the proportions each type of evidence contributed to teachers’ OTJs.

Student voice was weighted least heavily by junior teachers (9%), slightly more by middle teachers (13%), and more heavily by senior teachers (19%).

Observations contributed a very high proportion to all teachers’ OTJs, and were most heavily weighted by junior teachers (71%), with the proportion dropping for middle teachers (63%) and slightly less again for senior teachers (59%).

Within the observations section, running records were analysed as a subsection as these constituted a significant proportion. Running records were weighted most heavily by middle teachers (32%), then senior teachers (29%), and were weighted least heavily by junior teachers (22%).

Norm-referenced, standardised tests were moderately weighted, contributing 14% to junior teachers’ OTJs and 24% to middle teachers’ OTJs. These tests were weighted slightly less by senior teachers (22%) than middle teachers.

### 4.1.4 Sources of professional learning accessed to make reading OTJs

Participants were asked to select the sources of professional learning and information they had accessed in order to make OTJs of student achievement in relation to the reading National Standards; and to rate the usefulness of these sources.
Table 3

Sources of Professional Learning, Rated by Usefulness (n=34)

<table>
<thead>
<tr>
<th>Source</th>
<th>Yes I used this</th>
<th>Not useful</th>
<th>Moderately useful</th>
<th>Very useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education Gazette</td>
<td>13</td>
<td>0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Existing teacher knowledge</td>
<td>31</td>
<td>1</td>
<td>9</td>
<td>17</td>
</tr>
<tr>
<td>New Zealand Curriculum Reading and Writing Standards</td>
<td>34</td>
<td>0</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Other Ministry of Education publications</td>
<td>27</td>
<td>0</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Professional development workshops</td>
<td>24</td>
<td>1</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>TKI</td>
<td>18</td>
<td>1</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Thirty-four participants answered the first part of this question by selecting the sources they had used, but not all participants rated each of these sources.

All 34 teachers who answered this question used The New Zealand Curriculum: Reading and Writing Standards for Years 1-8 as a source of information. Other widely used sources of information included existing teacher knowledge (91%), other Ministry of Education publications on reading (79%), and professional development workshops (71%). TKI was accessed by just over half of the participants and most of those who rated this source found it moderately useful.

Few teachers found any of the sources not useful. The most useful source was existing teacher knowledge, which 63% (of participants who rated this item) reported to be very useful.

4.1.5 Additional support identified for development in making reading OTJs

Participants were asked to select the kinds of additional support they would find useful for future learning in making OTJs in relation to the reading National Standards, and to rate the usefulness of these.
Table 4

*Types of Additional Support, Rated by Usefulness (n=34)*

<table>
<thead>
<tr>
<th>Types of Additional Support</th>
<th>Not useful</th>
<th>Moderately useful</th>
<th>Very useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>National tests of reading</td>
<td>6</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Online exemplars</td>
<td>0</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Professional development workshops</td>
<td>0</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>Updated Ministry of Education publications on reading</td>
<td>0</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Thirty-four participants answered some part of this question, but did not necessarily rate every style of support. Eighty-five percent of participants thought further professional development workshops would be very useful. Other very useful forms of further professional development were online exemplars (53%) and updated Ministry of Education publications on reading (56%). When the moderately and very useful categories were combined, 100% of participants rating online exemplars, professional development workshops, updated Ministry of Education publications on reading and other (specified as talking to colleagues and other schools on two questionnaires) thought these resources would be useful.

Opinion was divided over the usefulness of national tests of reading. Of the 25 participants rating this item, 18% thought these would not be useful, 26% thought these would be moderately useful, and 29% thought these would be very useful. When the moderately and very useful categories for national tests were combined, 76% of participants rating this item thought national tests would be useful.

4.1.6 Participants’ reported confidence in reading OTJs

Participants were asked to rate their confidence when making OTJs of student achievement in relation to the reading National Standards. Confidence was analysed by year and experience level.
Confidence When Making Reading OTJs by Year Group Level Taught (n=34)

<table>
<thead>
<tr>
<th>Year Group Level Taught</th>
<th>Not at all confident</th>
<th>Not very confident</th>
<th>Quite confident</th>
<th>Confident</th>
<th>Very confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior (n=13)</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Middle (n=13)</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Senior (n=8)</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>All teachers</td>
<td>0</td>
<td>4</td>
<td>17</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 6

Confidence When Making OTJs by Experience of Teacher (n=34)

<table>
<thead>
<tr>
<th>Experience of Teacher</th>
<th>Not at all confident</th>
<th>Not very confident</th>
<th>Quite confident</th>
<th>Confident</th>
<th>Very confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice (n=7)</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Experienced (n=15)</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Extremely experienced (n=12)</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>All teachers</td>
<td>0</td>
<td>4</td>
<td>17</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

No teachers reported that they were not at all confident, and 12% reported that they were not very confident when making reading OTJs. Most (79%) participants reported that they were quite confident or confident when making OTJs in relation to the reading National Standard. A further 9% were very confident in making reading OTJs.

Forty-six percent of junior and 54% of middle teachers rated their confidence reasonably highly (in the confident and very confident categories), whereas no senior teachers rated themselves in this group. Senior teachers had the least spread of ratings, confined to quite confident (75%), and not very confident (25%).

Table 3 shows that participants cited existing teacher knowledge as their most useful source of information when making reading OTJs, indicating that the experience level of the teacher might impact on their confidence in making OTJs. Novice teachers were most likely to select not very confident, with 43% of novice teachers selecting this rating. Experienced teachers were the most
confident group, with 40% of this group selecting confident or very confident. Extremely experienced teachers were a little less confident, with 33% of this group selecting confident or very confident. The sizes of each of these sub groups of participants was small (n=7 to n=15) making it easy for the results to be skewed.

4.2 Focus groups

All literacy leaders who participated in the focus groups were classroom teachers, reporting on both their own practice and their schools’ practices in making reading OTJs. Discussion in the focus groups centred on five areas, which are identified in the following sections.

4.2.1 Use of information gained from assessment tools and processes

The literacy leaders discussed the information about students’ reading competencies they gained through using the assessment tools and processes listed in the questionnaire. Much of the discussion centred on formative uses of the tools, with teachers directly relating the information gained from an assessment to their practice in the classroom, e.g., using Burt to gain information about students’ decoding, using school entry tests to offer differentiated instruction, and using conversations to select appropriate texts.

Literacy leader four (LL 4) described how conversing with and observing students enabled her to gain information about students’ understanding of a text, and to adapt her teaching to ensure all students are participating.

I observe what they are reading as their personal choices and try to read them a bit more if they’re not making challenging choices… it’s to stop all the time and recap: “what does that mean and what have you got out of that?... Observing which ones are telling me and which ones are sitting back (LL4).

Three literacy leaders discussed how - at a school-wide management level - they used data from norm-referenced, standardised assessments to map trends from year to year, to identify target students, set goals, and measure progress.
4.2.2 Advantages and challenges when using multiple sources of evidence

All literacy leaders commented on the value of using several tools to give a fair and balanced assessment of students’ reading competencies, with two literacy leaders pointing out that – while the term OTJ was new – they had always used multiple sources of evidence to assess students.

I think there are real advantages... I think it’s broadening our understanding of a child’s reading. We’re making a richer assessment, we’re asking for a richer response from children as opposed to a running record (LL 6).

You’ve got to take everything. Some kids can’t sit tests and it’s not giving you a full picture of them. Some kids freeze up if they have to read out loud to you because they think it’s all about fluency and then they don’t listen to what they’re saying. So you need the whole mesh of everything to actually give the picture (LL 4).

Two literacy leaders shared how the National Standards had helped to shift their practice of depending on running records to assess junior reading, and had enabled them to think more broadly about junior reading.

Another literacy leader described how using multiple sources of evidence “confirms the predictions you’re making about that child,” and described how conflicting pieces of evidence “makes you rethink about what you’ve thought about that child” (LL1).

The literacy leaders identified challenges in using several sources of evidence to make an OTJ. Making consistent judgements in relation to National Standards using several pieces of evidence was an issue raised by all literacy leaders. Five literacy leaders identified inconsistent results between different tools as a challenge, for example,

Sometimes it’s very hard to see a pattern in the sources you are using (LL1).

They don’t match up... you’ll do a standardised test and you’ll get ‘ok this is where this child fits’ and then you do a different test and they’re fitting somewhere else and then [someone else] will do something with them and they fit somewhere else (LL4).

Four teachers identified getting evidence of students’ reading behaviours as a challenge and two literacy leaders identified time as an issue in working with several sources of evidence. One stated:

I do think we have to address the workload, we just need to as far as a busy teacher’s life actually goes and what we can fit in when we’re making OTJs in writing and maths and everything else. So I really support that – please can we have some direction and professional development in doing that? LL6.
The literacy leaders were clear that they needed both evidence gathering and judgement making to be efficient processes in order for them to be sustainable.

4.2.3 Weighting the importance of assessment tools and processes when making junior, middle and senior reading OTJs

Literacy leaders were asked to discuss which tools and processes they thought should hold most weight when making junior, middle and senior OTJs. Their weightings were different to those of the questionnaire participants.

The participants described how changes in reading development meant that different assessment tools and processes were needed to assess junior, middle and senior students: “I suppose.. the junior classes you’re teaching them to read whereas you’re moving into reading to learn by the time you get to year four” (LL 5). When discussing making an OTJ for junior students, three literacy leaders thought that running records should hold the most weight, and one literacy leader thought the six year net should hold the most weight.

When discussing appropriate assessment tools and processes for middle students, one literacy leader discussed the value of using a standardised test in the middle school to gather information about comprehension and vocabulary development and four literacy leaders recommended a balanced approach in teaching and assessing reading; as opposed to the decoding focus recommended for junior students. They discussed the importance of observations to gather information about comprehension strategies. LL 6 described her reasoning:

I’m right at the beginning of the middle school so I’m thinking are these [year four students] really comprehending well so they can use their reading for learning? I’m going to be bridging that gap making sure they’re ready for later middle school by understanding their comprehension strategies (LL6).

When discussing which tool or process would hold the most weight when making an OTJ for senior students, three literacy leaders said the PAT: Reading would hold the most weight, and one literacy leader said that observing student’s thinking skills as they read across the curriculum would hold the most weight for senior students. LL 1 described why running records were used with less frequency in the senior area of their school: “running record [s are] a lot to do with the strategies they’re using to decode... it’s not going to test comprehension.”
4.2.4 Professional learning and development

The literacy leaders agreed that they felt underprepared for implementing the reading National Standards. One literacy leader described her experience in getting started using the National Standards document: “We were just left on our own, like, here’s the book – go for it!” (LL5)

While *The New Zealand Curriculum National Standards in Reading and Writing for Years 1-8* (Ministry of Education, 2009) was the primary implementation resource available free to all teachers, some literacy leaders identified challenges in using the document. LL4 noted that the standards for senior students were very broad and open to interpretation: “We find that you could really make any child fit it. They’re not narrow enough to weed children out.” LL5 raised the concern that schools may use narrow aspects of the early reading Standards:

> Down at the lower levels the standard is green but it’s 90% seen so you can... make all sorts of children fit into that, children that aren’t actually reading. Because if it’s seen you can get them to remember a text just about and tick them – *they’ve passed!* So I think that’s actually quite a fall down because we’re making children fit the standard that aren’t actually at standard (LL5).

One literacy leader described the reclassification of *School Journal* as a helpful change.

Most literacy leaders had attended workshops or literacy leaders’ meetings based around implementing the reading and writing Standards. LL3 described her frustration at attending workshops that did not address her needs: “Every course I’ve been to that is literacy, we put our hand up and say ‘can we do moderating in reading please?’ And we never do it. Because nobody seems to have the answers” (LL3).

Five of the six literacy leaders expressed a need for more professional development in making OTJs with relation to the reading National Standards. They felt the need was particularly strong in the area of gaining consistency within and between schools, and that specific training in the area of moderation was required. Their strong preference was for face-to-face, practical workshops.

Two literacy leaders highlighted the need for training all members of staff, not just the literacy leader from each school:

> I think more workshop type things. But I think it’s really important it’s just not the literacy leader that goes out and comes back and feeds back. Because everybody when they go to professional development; they all come back with something different. I think every teacher needs to be targeted (LL1).
By taking it through the literacy leader you’re diluting it to a degree, rather than everybody getting the experts’ opinion on it (LL2).

TKI was accessed by a couple of literacy leaders, but finding time to browse the site, and knowing what was available online to search for, were issues discussed by two literacy leaders: “We’ve got a lot of staff who use TKI on a regular basis and find that helps. But you see a lot of it is word of mouth have you seen this site? – And who is going to provide that word of mouth?” (LL1). Another literacy leader found that online exemplars were helpful when working with the curriculum, but less helpful for the Standards.

4.2.5 Consistency and moderation of OTJs

The level of confidence reported in the questionnaire data was not so evident in discussions between literacy leaders, who displayed limited confidence in teachers’ ability to achieve consistency of reading OTJs. For example,

How do you moderate reading? That’s a big challenge! (LL1).

How do you know what you are doing in your school is consistent – well it won’t be – with what’s happening with every other school in New Zealand? (LL4).

Yes, because what I think looks like a good reader at this level – you might have totally different ideas at your school (LL5).

Two literacy leaders shared their concerns that inconsistency of OTJs would result in the implementation of national tests:

But then, we don’t want to say that [consistency is an issue] and then have the government say ok you’ve all got a test to teach to. Because then the whole education system falls apart (LL4).

So we’re better to go along with this and make it fit and pretend we can do it (LL5).

Literacy leaders were clear that, although they were not yet confident moderators of reading OTJs, they were working on the goal of consistency in their classes and schools. Two literacy leaders described their process of becoming more accountable with their observations of student learning:

Personally for myself it’s backing it up with evidence. That is the growth for me. It’s no good saying just this, it’s actually show me! Where’s the evidence? And I would say that’s probably across our school (LL1).
Yes I think we’ve got to get a lot better and I guess we’re starting to - at actually recording everything that you do so we can go back and look at it. Maybe in the past we’ve just relied on doing it and remembering it and haven’t actually recorded it. But in recording it you can be a lot more specific and you can go back and look at that specific detail (LL2).

Most literacy leaders identified moderation as their next step in working on consistency of reading OTJs within their schools. Three literacy leaders described the creation of new documents or matrices to support their staff in making OTJs. When discussing the creation of matrices, a literacy leader cautioned that: “If each school does that it’s going to be a lengthy doubling up process” (LL2).

Most literacy leaders identified time as a challenge in exploring between-schools moderation, and agreed that within-school moderation was a more appropriate goal.

**4.2.6 New learning gained through making OTJs of students’ reading**

While participants valued their teaching experience when making OTJs for the reading National Standards, several literacy leaders described how their theories about reading had shifted as a result of assessing to the National Standards; suggesting that contemporary learning and the ability to reflect and change one’s practice is also of value when making reading OTJs. Most focus group members shared some aspect of new learning they had gained through making reading OTJs. Two described the impact the new policy had had on their teaching:

Well I think the new learning for me has been it’s a much bigger picture than one running record, it’s just huge. And I think for children who have perhaps been good word barkers in the past – we have said they’re doing well in reading, and perhaps they have made accelerated progress. The child who has struggled a little bit with decoding but has this wonderful inferential understanding, we have done a huge disservice to (LL3).

I feel like reading is at an exciting change point for us in the classroom because of the type of evidence we’re being asked to show. And it is so strongly based on… a creative response to comprehension. So the kids are working with the reading and they have to produce something and that becomes part of the evidence alongside the asTTle and a running record at the junior level (LL 6).
4.3 Summary of results

Questionnaire participants and literacy leaders participating in the focus groups reported using several sources of evidence when making an OTJ, but differed when weighting tools for junior, middle and senior levels. The quantitative and qualitative data shared similar patterns of professional learning accessed, and preferred for future development. Questionnaire participants were generally confident in their own reading OTJs, but literacy leaders were unconfident in the consistency of OTJs within and between classes in their schools. While literacy leaders were exploring ways to develop moderation and gain consistency of reading OTJs within schools, they felt that consistency between schools was at present an unrealistic goal.
Chapter 5 Discussion

This discussion (Chapter five) considers the quantitative and qualitative data reported in Chapter four with relation to other studies of National Standards implementation. Positive aspects and challenges in making reading OTJs are identified, and the implications, recommendations and limitations of this study are discussed at the end of this chapter.

5.1 Assessments and processes used when making OTJs in reading

Participants reported using multiple sources of assessment evidence, and triangulating data between the three main areas of an OTJ - conversations with students, observations of students reading, and assessment tools - when assessing in relation to the reading National Standards. Some of the sources of evidence used differed between teachers of different levels.

5.1.1 Use of conversations

Almost all participants reported using student voice when making an OTJ in reading, which is in accordance with The New Zealand Curriculum (Ministry of Education, 2007) and the National Standards. When investigating National Standards in 2010 (Education Review Office, 2010) and in 2012 (Education Review Office, 2012a), ERO found that using student voice to improve learning outcomes for students was a problematic area for New Zealand schools. ERO found that:

A key issue for teachers was finding ways to share information with students about their progress and achievement against the standards that enabled them to take an active role in assessing and reflecting on their learning and determining their next steps to meet or exceed the standard (p. 11).

An additional challenge identified by literacy leaders in the focus groups was the difficulty of gathering evidence of conversations with students, particularly with junior students who are less likely to have the writing skills to record their own goals and progress. This is of concern as it is difficult to moderate OTJs without some representative evidence to support teacher judgement. Participants in this study weighted student voice as contributing between nine and 19% to their reading OTJs, and three literacy leaders referred to its importance in the focus group interviews. However ERO’s findings (ibid.) suggest that the student voice used to inform teachers’ OTJs may not be specific and targeted enough to enable student involvement and to drive future learning. The findings of this study suggest that
participants are using student voice, but may need more supportive exemplars to use it purposefully and to gather evidence of representative conversations.

5.1.2 Use of observations

Observations of student reading - a broad category encompassing classwork, formative assessment tasks, and observing students reading in guided reading time and across the curriculum - were used by all participants and contributed a significant proportion to their OTJs (combined mean = 66%). This is a similar pattern to that found in Thomas and Ward’s (2011) study in which 86% of participants reported that specific classroom observations were of high importance when making an OTJ in reading. They also found that 75% of the specific classroom observations teachers reported using to support their reading OTJ were of a general nature (p. 14).

The current study did not gather and evaluate teachers’ evidence so it is not known how specific and targeted the observations reported in the quantitative data were. The literacy leaders interviewed indicated that while some participants were finding ways to provide evidence of students’ reading behaviours and contributions when reading; others were having difficulty in finding ways to capture evidence of conversations and observations, and highlighted this as an area where PLD was needed. An implication for professional development providers and literacy advisors is that teachers may require clear examples of what constitutes evidence of students’ reading behaviours, strategies, and processes in the classroom; and also examples of a variety of ways in which to capture this evidence.

5.1.3 Use of running records

All participants teaching junior students used a seen text running record when making an OTJ in reading for students after one, two and three years at school. This practice is informed by Ministry of Education documents (Ministry of Education, 2003, 2009), and is supported by research evidence – for example, Nicholson’s (2010) findings that running records are a valid measure of junior students’ reading, as their results correlate highly with other measures of early reading. Junior teachers reported that running records contributed a mean of 22% to their reading OTJs. While running records are developmentally appropriate for early readers, their usefulness in gaining information about fluent readers is very limited. Yet participants reported universal use of running records to assess middle and senior students’ reading, and attributed relatively high value to running records when making OTJs for the reading National Standards (middle mean =32%, senior mean = 29%). Similarly, 73% of Thomas and Ward’s (2011) participants (n=66 teachers of year one to eight students) rated running records as being of high importance when making reading OTJs (p.13). These are very large proportions to attribute to a procedure designed to observe young students’ early oral reading behaviours. Effective
Literacy Practice in Years 5-8 (Ministry of Education, 2006) suggests that teachers of middle and senior students may use running records as a formative tool when students are not making expected progress. While this scenario may account for some instances of participants’ use of the tool, the high weightings suggest that widespread use across all students in classes may be more likely. In addition to concerns about the validity of using running records to assess middle and senior reading, is the time-consuming administration that may impact on the class reading programme (Timperley H., 2004).

Literacy advisors and PLD providers need to find ways to shift school and teacher practice when information about appropriate use of running records has already been disseminated in many Ministry of Education publications on reading (e.g. Ministry of Education, 2000, 2006). Literacy leaders’ data indicated that some schools are working to shift their heavy reliance on the results of running records to provide summative assessment data. Over half of the literacy leaders interviewed described how National Standards had helped to broaden their literacy assessment and promote the use of additional sources of information. However the questionnaire data suggests that a dependence on running records to assess reading across all levels remains. A possible reason for this is that teachers may be more experienced at, and feel more comfortable when taking running records than when gathering observations or interpreting standardised tests. This would be a cause for concern, and imply that support is needed for teachers to develop skills in administering, interpreting, and using the information from a wider range of assessments.

5.1.4 Use of norm-referenced, standardised assessments

Ninety-four percent of participants in this study reported using one or more norm-referenced, standardised assessment tool when making reading OTJs. Dingle and Parr (2010) state that “the use of nationally standardised tests allows teachers and leaders to compare the level attained, or the progress they have measured in their students, with the expected level or average gain for that group of students” (p. 137). The use of at least one norm-referenced tool was a criterion for ERO when evaluating whether teachers were working well with the standards in 2012 (Education Review Office, 2012a). That nearly all participants meet this criterion is laudable, and in addition, their reported use of different norm-referenced, standardised tools was generally age appropriate: junior teachers were most likely to use the Observation Survey and/or the Burt word test, and middle and senior teachers were most likely to use the PAT reading test. In some cases unusual test use was reported (e.g., use of the Observation Survey with middle and senior students and use of PAT with junior students). It is possible that assessment of well below and above students may account for this, and in such cases the use of these tools may be appropriate for the developmental level of the student.
asTTle was very infrequently used by participants in this study, with just two teachers of middle school students reporting that they used the tool. This was a surprising finding that may be particular to this geographic region, or alternatively, one that may apply to other rural areas. asTTle was designed to provide next steps for student learning and to support teacher-student dialogue about assessment, and these features make it a useful tool when gathering evidence for OTJs. It is possible that the rural area the sample was drawn from has limited teacher access to professional development of asTTle. In addition, the ICT constraints of small schools, and the challenges accessing reliable internet in remote areas may account for low asTTle use as it requires access to the internet and enough computers for students to sit the assessments in a reasonable timeframe.

The reported use of norm-referenced, standardised tools by participants in this study does not support the claim that National Standards has led to more formal testing of students. A concern raised by the New Zealand Education Institute (union for primary teachers) (Thrupp & Easter, 2012) was that National Standards would lead to more testing of students, and eventually build a culture of high stakes testing as has occurred in other countries. New Zealand’s National Standards have no testing regime attached to them, and decisions about assessing students are left up to individual schools. While nearly all participants reported using norm-referenced, standardised tests, the data from these was moderately weighted, and contributed less than a quarter to their OTJs of student reading. Earlier studies e.g., (Thomas and Ward, 2011; Wylie & Hodgen, 2010), similarly report teachers using a balanced range of evidence to make their OTJs, and one could tentatively conclude that National Standards have not resulted in an increase of formal testing of students.

5.1.5 Challenges in using existing norm-referenced, standardised assessments

A challenge identified by literacy leaders was the difficulty of making an OTJ with pieces of assessment evidence that gave conflicting results. The participants in Poskitt and Mitchell’s (2012) study also faced a similar predicament when synthesising “information from formal and informal assessments, especially when there was some divergence in the data” (p.68). A possible explanation is that existing formal, standardised assessment tools used by teachers when making OTJs were not designed to measure the National Standards. When asked whether national tests of reading would support participants to make their reading OTJs, opinion was divided. Many participants (76% of the 25 participants who rated this item) thought that national tests would be a moderately or very useful form of additional support when making reading OTJs, with the remaining 24% rating national tests as not useful. Two focus group members shared their concern that if schools were vocal about the challenges they were facing in gaining consistent OTJs; national testing (that is, nation-wide testing using a universal standardised assessment) would be implemented. The quantitative data suggests that
many other participants were interested in having a specific test to support their reading OTJs; possibly because of the interpretation needed when using existing tests with the Standards.

A section on TKI entitled *Alignment of Assessment tools with National Standards* (Ministry of Education, 2010c) suggests that standardised test results using a bell curve of normed achievement need differing interpretation to meet the aspirational nature of the reading Standards:

> Test norms are based on what the average student of a given age *can* do; the standard relates to what all students *should be able* to do, if they are on track for a successful educational outcome. In some areas of learning, the two coincide – the average student is at the required level. In other areas, there may be a general shortfall – only high-performing students are likely to reach the standard, with others needing to improve their achievement in order to do so (ibid).

The Ministry of Education is engaged in on-going research evaluating National Standards’ achievement expectations with relation to the Observation Survey, STAR, PAT, and asTTle. In each case sample student achievement that would have been considered average in the past is rated as ‘below standard’ (ibid). While TKI clearly states that “This material is provisional and under review, Overall Teacher Judgments should be supported by a range of evidence, as different assessment methods look at different aspects of student work” (ibid.); the need for work in this area suggests that current standardised assessment tools may not be easily used when making OTJs. It seems conflicting to ask teachers to continue to use existing tools, yet interpret the tests’ results differently to their developers’ intent. It may also impact negatively on teachers’ assessment literacy (that is, their knowledge and understanding of assessment processes and uses) by devaluing the research provided in the test manuals for PAT and STAR. In 2011 ERO recommended further guidance from the Ministry regarding alignment of assessment tools with the Standards (Education Review Office, 2011, p. 23), yet in 2012, ERO reported that schools were still struggling to align existing evidence with the Standards (Education Review Office, 2012a, p. 17).

The limited ability of norm referenced, standardised assessments to measure the aspirational Standards may undermine the validity of comparison studies intended to provide a check on OTJs. An early study (Thomas & Ward, 2011) found that teachers’ judgements of students considered well below, below, at, and above correlated with assessments gathered using other data gathered by the Ministry of Education (p.27). More recently the Minister for Education Hekia Parata (Hartevelt, 2012) compared PIRLS (Progress in International Reading Literacy Study) reading results with National Standards’ reading data, finding similar patterns of student achievement from the sources. However the PIRLS assessments measure what the average 4th grader *can do*, and provide norm-referenced assessment data based on reading material “suitable for fourth grade students in content, interest, and reading ability” (Mullis & Martin, 2010 p12). In contrast, New Zealand’s National Standards measure what students
should be able to do. As the two constructs are different, comparisons between data sets measuring the different constructs must be viewed with caution. A new initiative may provide more reliable data: The National Monitoring Study of Student Achievement (a partnership between NZCER and the Educational Assessment Research Unit – University of Otago) is a new project that will assess the achievement of a representative sample of New Zealand students. Under contract to the Ministry of Education, the project may help to revise some aspects of the National Standards that prove too challenging for most students to attain.

5.1.6 Triangulating evidence to make an OTJ

Just over a third of the questionnaire participants had difficulty when reporting on the weightings of each assessment tool and the process used when making a reading OTJ (the design of this question is addressed in the limitations at the end of this chapter). In general, the remaining nearly two-thirds of the questionnaire participants reported that observation and informal assessment contributed a large proportion to their OTJs, and norm-referenced, standardised assessment and student voice a lesser proportion. These proportions were different to those reported by the literacy leaders in the focus groups, the majority of whom thought running records should carry the most weight in junior OTJs, observation of comprehension processes should carry the most weight in middle OTJs, and the PAT: Reading test should carry the most weight in senior OTJs. It is possible that the difference in quantitative and qualitative findings arose because questionnaire participants reported on how they made OTJs, whereas literacy leaders reported on how they thought they should make OTJs. It is also possible that the questionnaires were answered quickly, whereas the focus group questions were answered in a collaborative setting that enabled participants to consider and clarify their responses. An alternate explanation - supported by Poskitt and Mitchell’s (2012) findings that teachers’ understandings of how to derive OTJs were varied and inconsistent (p.66) - may be that a shared understanding within and between schools about making consistent OTJs at different levels has not yet been reached.

5.2 Professional learning and development

5.2.1 Use of professional learning and development resources in print and online

All participants reported using The New Zealand Curriculum Reading and Writing Standards for Years 1-8 (Ministry of Education, 2009) as a source of professional learning when making reading OTJs; a much higher proportion than reported in an earlier study which found a quarter of participants were
using the document to inform their reading OTJs (Thomas & Ward, 2011). All participants who rated the usefulness of the document reported it to be moderately or very useful, though two literacy leaders in the focus groups identified aspects of the document that were confusing to work with.

Another source of freely accessible information for teachers who are seeking to interpret or clarify the reading Standards is TKI – the online community for New Zealand teachers. Making OTJs, moderating, using appropriate assessment tools and processes, and mapping these existing tools appropriately onto the National Standards are topics and processes explored and described in articles on TKI, not in print. Fifty-three percent of participants reported using TKI, and 44% rated its usefulness as a source of professional learning when making reading OTJs. Most participants rated TKI as a moderately useful source, but when literacy leaders were asked how useful TKI was as a tool for learning about National Standards, one said it was a challenge finding out what was on the site, and others said they used the online community to access teaching resources rather than as a source of professional learning. The effectiveness of transmitting information using a network of web pages and uploaded files is uncertain. There are many resources on TKI relating to the reading National Standards but the time needed to read and process these resources is likely to be challenging for a teacher, literacy leader or principal who is dealing with the demands of the wider curriculum at the same time. Very high numbers of participants (100% of the 32 rating this item) thought that online exemplars would be moderately or very helpful when making reading OTJs. It is possible that interactive exemplars and film clips of teacher demonstrations may be of more practical use than online documents. An implication for the Ministry of Education is to think more widely about ways to provide professional resources to all teachers. High numbers of participants (100% of the 29 rating this item) rated updated Ministry of Education publications on reading as moderately or very helpful. This implies that print resources may be a preferred option for teachers’ professional learning as many teachers may not be accessing TKI. No known research is available investigating the effectiveness of using TKI for professional learning and development.

5.2.2 Use of professional learning and development workshops and courses

Seventy percent of participants attended PLD courses or workshops when learning to make reading OTJs and of those who rated the quality of this PLD, nearly all found it moderately or very useful. This is more positive than some of the PLD experiences of teachers reported in earlier studies (e.g., Education Review Office, 2010; Cameron, 2010; Wylie & Hodgen, 2011) which found that many teachers experienced the PLD they had accessed as unhelpful, confusing or inadequate. While the participants in this study were generally positive about their experience with PLD, 100% (of participants answering this question n=34) reported that they would find further professional
development workshops useful. Though literacy leaders accessed PLD courses, they described feeling underprepared to use the Standards, suggesting that initial training in National Standards implementation was useful but not sufficient. The efficacy of accessing timely quality PLD is highlighted in ERO’s (Education Review Office, 2012b) recent report: *Working with National Standards: Good Practice*, in which the case study schools described how engagement in PLD over time helped to build teachers’ assessment capabilities and develop consistent practices. A difficulty the participants in this study may face is on-going access to quality PLD, as ERO’s (Education Review Office, 2012a) recent report found on-going difficulties for small and rural schools in accessing quality PLD.

Location and quality of PLD are not the only challenges facing some teachers in learning how to implement the reading Standards. Opportunities to attend courses and workshops, and to participate in professional discussion are dramatically reduced in schools where management opposes National Standards, or where professional development in this area is not prioritised. Wylie and Hodgen (2010) suggest that early opposition to the standards may have left some teachers without access to the professional development they need in order to understand and use the National Standards, as they say: “The New Zealand self-managing schools context, where schools make their own decisions about how they will give effect to national policies, has provided uneven ground for the introduction of the National Standards” (p. 18).

While New Zealand’s National Standards emphasise teacher professionalism and empowerment to make decisions suited to each school’s context; instituting new and challenging policies, without free, clear and consistent professional development for all teachers is likely to prove overwhelming rather than empowering.

### 5.3 Confidence, consistency and moderation

The confidence participants reporting feeling when making their reading OTJs did not translate to confidence that reading OTJs across schools and clusters would be consistent. In 2010 Thomas and Ward (2011) found that 99% of teachers were confident in the accuracy of their reading OTJs. In the quantitative data for this study, 79% of participants reported feeling moderate to high levels of confidence when making reading OTJs. It is possible that the lower confidence level reported in this study is an outcome of the limited PLD reported earlier. An alternate argument is that as teachers have had more time working with the complexities of the reading Standards, confidence in one’s OTJs may have dropped rather than increased. ERO (Education Review Office, 2012a) included two criteria (p.24) based on teacher confidence when reviewing 439 schools for the report *Working with National Standards,*...
Standards to Promote Students’ Progress and Achievement: The first was: “Teachers are confident in making overall teacher judgements based on multiple sources of evidence which, where possible, include at least one norm-referenced tool.” Based on participants’ reported data, it seems likely that most would meet this criterion, hence the moderate degree of confidence when making OTJs.

Now schools are underway with gathering evidence and making reading OTJs, the next challenge is to achieve greater consistency of these OTJs through moderation. The literacy leaders shared universal concern that reading OTJs were inconsistent within and between their schools. Poskitt and Mitchell describe the underlying difficulties contributing to a lack of consistency between judgements:

In conclusion, in June 2010 teacher judgments across the ten schools were problematic because there was not common understanding of what constituted OTJs, understandings were rarely supported by clear criteria and exemplars, and few schools had robust moderation processes in place. (Poskitt & Mitchell, 2010, p. 11)

Two years later, the participants in this study reported a shared understanding of triangulating data when making OTJs and were using mostly appropriate assessment tools for their teaching level. Literacy leaders reported that they were working towards consistency in various ways: through gaining a shared understanding of suitable assessment tools in staff meetings and discussions, creating matrices and updating school assessment, and aiming for better consistency within their schools before engaging in external moderation. They also acknowledged that it takes time to get teachers working in the same way. Poskitt and Mitchell (2012) describe how “engaging in assessment moderation appears to help teachers resolve formative and summative assessment tensions and strengthen links between pedagogy, curriculum and assessment” (p.71). The phrasing of ERO’s (Education Review Office, 2012a) second criterion on OTJs: “Teachers are continuing to develop confidence in moderating their data” (p. 24) implies that achieving consistency through moderation is an on-going process and not one that teachers are expected to have perfected. Although they were not yet engaged in formal moderation of reading OTJs at the time of the focus group interviews, the literacy leaders were preparing the groundwork for doing so. Poskitt and Mitchell argue for the importance of engaging in the actual moderation stage, claiming that this will provide the best learning and understanding for teachers:

Although having clarity about the composition of OTJs and standards, appropriate evidence to underpin those judgements and using exemplars and verbal descriptors to support the judgements is likely to lead to greater consistency (Sadler, 1998; Gardiner et al., 2000), it is only through moderation processes that teachers will reach deeper understanding of standards (Wyatt-Smith et al., 2010, cited in Poskitt and Mitchell, 2012, pp.72-3).
ERO consistently cites moderation as a high priority in working effectively with the Standards. ERO’s examples of good practice found that all of their case study schools were engaged in moderation processes, and had been supported by PLD to do this. But to return to an earlier point, the resourcing, distribution and quality of PLD in New Zealand appears to be uneven, with very small and remote schools particularly disadvantaged. A concern raised by literacy leaders was that there was not adequate PLD available to support their schools to moderate reading. The other jurisdiction using teacher judgements - England and Northern Ireland have attempted to achieve consistency through tight guidelines for the teacher judgements in assessing students at the end of year six, with a mandated requirement for teacher and school moderation, and for teachers to submit their judgements to the Department for Education before students sit their National Curriculum tests (Department for Education: Standards and Testing Agency, 2012). In New Zealand an alternate method to improve consistency of teacher judgements is in development. The Progress And Consistency Tool (Forrest, 2012) is being developed to support teachers to make consistent and reliable judgements. Informal and formal assessment data are collated, weighted, and displayed in tandem by the online tool, allowing the teacher to consider the balance of their tacit knowledge and formative classroom assessments, with the data from their norm-referenced assessments. The PACT tool has the potential to support teachers to use valid tools for their level, and to show (through the descriptors provided) useful classroom evidence to look for. The tool is currently under trial and its effectiveness in supporting teachers to make reliable OTJs for the reading National Standards is not yet known.

While participants’ self-rated levels of confidence were reasonably high, consistency was the major challenge identified by literacy leaders, and their primary area of need for PLD. Many studies cite moderation as an important area for schools and teachers to develop, and the literacy leaders in this study were beginning work in this area.
Limitations

This research project was a small scale investigation into how teachers make their OTJs for the reading National Standards. It had limited external validity because of the small size of the sample, and its restricted geographic coverage. It is possible, for example, that schools in large urban areas may have access to more professional development opportunities and be involved in cluster groups for moderation. The small size of the sample meant that junior, middle and senior teachers were clustered together to look at trends in their assessment choices. Because of this, it was not possible to gather detailed information comparing assessment for the Standards at each year level.

The small size of the study also meant that the researcher chaired and transcribed focus group interviews and entered all data herself.

The study design meant that a snapshot of participants’ information on their reading OTJs was gathered mid-year, 2012. All data were self-reported by the participants and it was beyond the scope of this study to gather and evaluate examples of teachers’ OTJs of reading. Participation in the study was voluntary and it is possible that the schools and participants who chose to take part in the study were already confident in this subject area. The tools used to investigate participants’ OTJs of reading have some limitations. Questionnaires typically have a low return rate (for this study approximately 59%) and misunderstandings are unable to be corrected (Gillham, 2000, cited in Mutch & Limbrick, 2010, p.94). Cresswell (2009) describes how the focus group setting can advantage articulate members and disadvantage quieter ones.

The design of the questionnaire may have led to some confusing results. Calculating the percentage each assessment tool or process contributed to an OTJ appeared to be a challenging question for participants. It is not known whether the confusion in answering this question was a reflection of the challenge in weighting the different tools and processes when making an OTJ, or whether this question was unclearly worded.

Teachers of a given year group level may use assessment tools and processes more suited to older or younger students when assessing students who are well below or above the National Standard for their level. It is not known to what degree this practice may account for the high use of running records in middle and senior classes in the study.

Should the study be replicated, a wider geographic sample including urban areas, and a revision of the questionnaire to clarify weighting of tools, and to specifically focus on the achievement of students ‘at’ National Standard are recommended.
Recommendations

The findings of this study suggest that in order to ensure more reliable, valid and efficient OTJs in reading the following are required:

1. Updated informative media for teachers are needed which give clear, concise guidelines to:
   - outline appropriate assessment tools and processes to use when gathering evidence of observations, conversations, and assessment for each level of the reading Standards;
   - demonstrate a range of ways to gather and record evidence of students’ conversations and cognitive behaviours;
   - provide examples of teachers drawing multiple sources of evidence together and making an OTJ, both in circumstances when the evidence is consistent, and when it is not.

It is strongly recommended that guidelines are provided in print, as many participants in this study were unaware of pertinent resources available on TKI. A downloadable PDF may be a cost-effective way to ensure that all teachers can access consistent information. Over time, it is recommended that a cohesive handbook be developed to support schools and teachers to make reliable and consistent reading OTJs.

2. Professional learning and development for all teachers is needed. This should include:
   - free and consistent provision of professional development workshops to all teachers to ensure that all schools in New Zealand have the knowledge and skills to work successfully with the reading Standards; especially in the areas of gathering classroom evidence, gaining consistency of OTJs, and moderating OTJs;
   - a self-review matrix to enable teachers and schools to evaluate their progress toward working well with the reading National Standards, with the most competent stage including OTJs that were made with reliable assessment evidence gathered across the curriculum (Weal & Hinchco, 2010), and between-schools moderation of a sample of OTJs.

3. Investigation into the levelling of the Standards is needed to ensure that they are set at an achievable level for students and teachers. This could be achieved by:
   - the Ministry of Education using data from the National Monitoring Study of Student Assessment to investigate the reasonableness of aspirational aspects of the Standards. In cases where there proves to be a large difference between a standard and what an average student of
that age can do, it is recommended that that standard be revised and that all teachers are advised of the revision.

4. Further research is recommended to investigate the themes, trends and findings suggested by this small scale study:

- A larger scale study with a broader geographic and socioeconomic sample is recommended to determine whether the strengths and challenges in making reading OTJs reported by the participants in this study are common to other teachers in New Zealand. It is recommended that further research investigates more deeply into how teachers make their reading OTJs, by a) gathering examples of teachers’ conversations, observations and assessments and comparing them to evaluative criteria, and b) interviewing teachers about how they use this evidence, norm-referenced data, and the Standards as they make an OTJ.
Conclusions

The introduction of National Standards in New Zealand has required that teachers examine their assessment practices and, while retaining their competence in formative assessment, to now sum up a student’s achievement in relation to the relevant Standard. In reading, a lack of recent professional development and a reliance on tacit knowledge of reading development indicated that this may be an area in which teachers had particular difficulty. Previous studies have investigated National Standards implementation as a whole, or OTJs as a whole, with little differentiated data provided on the reading, writing and mathematics Standards. When the limited findings on reading OTJs were examined, little data was available about how OTJs were made at different year levels. This study endeavoured to explore how teachers of junior, middle, and senior students made their OTJs in reading; using self-reported data from questionnaires and focus group interviews.

This study revealed that teachers were using a range of appropriate data to make their reading OTJs. In most cases, the assessment tools and processes used to assess reading were appropriate for the level taught. Teachers referred to published documents to support their understanding of OTJs and literacy leaders used staff discussions, and updated policies and matrices to support teachers in working towards more consistent judgements. Participants were generally confident in their reading OTJs. Literacy leaders reported that within-school moderation was a next step for their schools and shared positive new learning from working with the Standards.

Of concern, however, was that teachers of middle and senior students appear to be relying on running records as an assessment tool although these may not provide useful information about their students’ reading. Other areas of concern were: that literacy leaders felt that they were underprepared to implement the standards, and were not confident that reading OTJs were consistent. Participants felt they needed more PLD, particularly in gathering evidence of conversations and observations, and in moderating reading. ERO suggests that teachers in rural areas such as this one may struggle to access quality PLD, and the researcher recommends that the Ministry of Education address this equity issue.

Resources on TKI did not appear to be well accessed by literacy leaders, and other methods of providing information should be explored. An additional concern was the apparent difficulties aspirational standards pose when using norm-referenced, standardised tools. New initiatives such as PACT and NMSSA may provide research to support either a simple alignment process between existing tools and the Standards, or an alteration in some criteria of the Standards. Future research is recommended to explore examples of teachers’ reading OTJs - particularly the way evidence of conversations and observations is gathered and used, and how norm-referenced data is used - to form OTJs of students’ reading in relation to the National Standards.
Appendix 1

Questionnaire

1. Which year level/s do you teach? (Please circle as many as apply to you).
   Year 0  Year 1  Year 2  Year 3  Year 4  Year 5  Year 6  Year 7  Year 8

2. How long have you been teaching in New Zealand primary schools? (Please circle the option that applies to you).
   0-2 years  3-5 years  5-10 years  10-20 years  20-30 years  Over 30 years

3a. Read the following options and select the assessment tools and processes you use when making an overall teacher judgement of student achievement in relation to the reading national standard.

3b. Estimate the percentage each assessment tool or process contributes to your overall teacher judgement of student reading achievement.

<table>
<thead>
<tr>
<th>Assessment resource bank</th>
<th>Yes I use this</th>
<th>It counts as x% of my OTJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>asTTle reading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BURT word test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversations with student about their reading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information on students’ recreational reading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observation during guided reading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observation of student reading across the curriculum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAT reading test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seen text running record</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Six year net</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAR reading test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher made assessment</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. Where did you find information to enable you to make overall teacher judgements of student achievement in relation to the reading National Standard? Please tick the sources you used, and rate their usefulness in the appropriate box.

<table>
<thead>
<tr>
<th>Source</th>
<th>Yes I used this</th>
<th>Not useful</th>
<th>Moderately useful</th>
<th>Very useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education Gazette</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Existing teacher knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Standards publication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Ministry of Education publications on reading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional development workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TKI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other—please specify</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. What kind of additional support would you find useful when making overall teacher judgements in relation to the reading National Standard? Please tick the appropriate box for each suggestion.

<table>
<thead>
<tr>
<th>Support</th>
<th>Not useful</th>
<th>Moderately useful</th>
<th>Very useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>National tests of reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online exemplars</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional development workshops</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Updated Ministry publications on literacy teaching (eg. Effective literacy practice)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. How confident do you feel when making overall teacher judgements of student achievement in relation to the reading National Standards? Please tick the appropriate box.

<table>
<thead>
<tr>
<th>Not at all confident</th>
<th>Not very confident</th>
<th>Quite confident</th>
<th>Confident</th>
<th>Very confident</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2

Focus group indicative questions

*Can you talk about the relative importance of data from these different areas (assessment tools and processes listed in the questionnaire)? Why might this change from junior to senior classes?

*What information about students’ reading competencies do these particular tools provide for you?

*What are the challenges in making an OTJ using several sources of evidence?

*What are the advantages in making an OTJ using several sources of evidence?

*Can you talk about moderating reading OTJs in your school?

*What would you look for in a valid, reliable OTJ for the reading National Standard?

*Can you talk about new learning you have gained in the process of making OTJs in reading?

*What are your next steps as a teacher in this area?

*What are your next steps as a school in this area?
Appendix 3.

Assessment tools referred to in this study

<table>
<thead>
<tr>
<th>Assessment name</th>
<th>Brief description</th>
<th>Intended level of student as recommended by developer (does not necessarily fit with the New Zealand Curriculum or National Standards)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Resource Bank (ARBs) tasks</td>
<td>Comprehension activities and questions based on school journal and other levelled articles.</td>
<td>Year 3-10</td>
</tr>
<tr>
<td>asTTle</td>
<td>Standardised, norm-referenced, customisable tests of comprehension and vocabulary.</td>
<td>Years 5-10</td>
</tr>
<tr>
<td>Burt word test</td>
<td>Norm-referenced test of word reading.</td>
<td>6-13 years</td>
</tr>
<tr>
<td>Observation Survey/ Six year net</td>
<td>Battery of norm-referenced tasks to measure alphabet and sound knowledge, sight word reading, early writing skills and print concepts.</td>
<td>5-7 years</td>
</tr>
<tr>
<td>Progressive Achievement Tests: Reading (PAT)</td>
<td>Standardised, norm-referenced tests of comprehension and vocabulary.</td>
<td>Years 4-10</td>
</tr>
<tr>
<td>PM Benchmark</td>
<td>Informal prose inventory in which unseen text running records are accompanied by oral retell and comprehension questions.</td>
<td>5-12 years</td>
</tr>
<tr>
<td>PROBE</td>
<td>Informal prose inventory in which unseen text running records are accompanied by oral retell and comprehension questions.</td>
<td>7 years-adult</td>
</tr>
<tr>
<td>PROSE</td>
<td>Informal prose inventory in which unseen text running records are accompanied by oral retell and comprehension questions.</td>
<td>6-15 years old</td>
</tr>
<tr>
<td>Supplementary Test of Achievement in Reading</td>
<td>Standardised, norm-referenced tests of comprehension and vocabulary.</td>
<td>Years 3-9</td>
</tr>
</tbody>
</table>
Reference List


