Social moderation: Assessment for teacher professional learning

Esther Mary Smaill

A thesis submitted for the degree of

Doctor of Philosophy

at the University of Otago,

Dunedin, New Zealand

2018
Abstract

This thesis addresses New Zealand’s unmet need for additional teacher professional learning opportunities in assessment. It reports on a yearlong, multi-case study involving the teachers at three New Zealand primary schools. This study investigated whether teachers, working within the National Standards context, could use their participation in social moderation to help the sector meet the need for professional learning in assessment. It asked: *how and what do teachers learn about assessment through their involvement in social moderation?* Drawing upon both social constructionism (Crotty, 1998) and Wenger’s (1998) social theory of learning, the study used observations, interviews, and a questionnaire to produce data about teachers’ experiences of social moderation. The analyses of these data, which utilised key elements of grounded theory (Charmaz, 2006), yielded important insights into how involvement in social moderation can strengthen teachers’ assessment capability. The study identified that teachers were able to use their participation in social moderation to improve their understandings of assessment for learning principles and practices. It also demonstrated that taking part in social moderation enabled the participating teachers to learn about factors that can affect the dependability of student assessment information. Moreover, it showed that teachers believed that involvement in social moderation had contributed positively to their assessment capability.

This study also found that the teachers at the three participating schools garnered qualitatively different learning opportunities from their experiences of social moderation. These differences were linked to a series of school-specific conditions. These conditions included the amount of time that schools committed to moderation, the types of moderation activity that teachers engaged in, and the nature of the rationale that teachers developed to sustain their involvement in moderation. These school-specific factors shaped the extent to which participation in social moderation enabled teachers to take part in the formation of a school-wide, assessment-focused community of practice. The findings from this study indicate that using involvement in social moderation to develop such a community of practice increases the likelihood that all teachers are afforded opportunities to learn about assessment. The recommendations that arise from this study are intended to help schools strengthen their social moderation processes in ways that should enable teachers to harness additional, assessment-focused professional learning opportunities.
Acknowledgements

Just as it takes a village to raise a child, it takes a community to complete a doctoral thesis. I would like to thank a number of people and organisations for their support.

This thesis could not have come into existence without the cooperation and assistance of the teachers and principals at the three participating schools. It was an absolute privilege to attend staff meetings at each of these schools and to learn alongside the teachers. Thank you to the teachers who completed the questionnaire and to the principals who agreed to be interviewed. I would also like to extend additional thanks to the group of teachers with whom I worked the most closely at each school. Thank you for allowing me to attend your syndicate meetings and for making the time to participate in interviews. He mihi nui rawa atu ki a koutou katoa.

I am exceptionally grateful to the University of Otago for awarding me a PhD scholarship, and to the Humanities Division and the College of Education for conference funding. I would also like to thank the Educational Assessment Research Unit (EARU) for both accommodating me and providing me with some additional conference funding during the first half of my candidature. Being a doctoral student can be a solitary endeavour, and during my time at EARU I really appreciated having the opportunity to work within such a stimulating and collegial environment.

To my supervisors, Alison Gilmore, Lisa Smith, and Alex Gunn, thank you all so very much for your ongoing guidance, support, and encouragement. I feel incredibly fortunate to have had such a wise and wonderful supervisory team. I cannot thank you enough for the rich, timely, and insightful feedback that you have provided me with throughout my doctoral journey. I am extremely privileged to have had the opportunity to work with and learn from each of you and I will miss our meetings.

I have a wonderful group of friends who have trodden this path by my side. Thank you for asking after my work and thank you, at times, for not asking after my work. Thank you for the coffees and the wines and the phone calls, emails, and texts. Thank you for running up mountains with me and thank you for having my kids around to play. In particular, I would like to thank Vanessa Manhire for so generously proofreading my final draft. Likewise, I am indebted to the wonderful Judy Robinson who has, for several years, gifted me child-free Tuesdays to help me complete this project.
Extra special thanks go, of course, to my amazing whānau. To my mother, Barbara – who has taken leave from work and flown from the opposite end of Aotearoa to look after my children so that I could work through the school holidays – thank you for all your love and care. You set the bar so high. Thank you for teaching me about the things that matter the most. To my father, Bruce, the first Dr. Smaill in our immediate family: you have a lot to answer for, Dad! Thank you for believing in me and encouraging me. Thank you for showing me what hard work looks like and for always being on the end of the phone when I have needed you. I couldn’t have asked for better or more loving parents. To my beautiful siblings, Anna and Chris, and to all of my wonderful “in-laws”: thank you for always asking after my thesis and for offering me endless support and encouragement.

To my incredible children, Nikau Awatea and Kahu Kura, who came along for the full ride, and Huia Pōtiki, whom we welcomed partway through this journey: thank you all for being so understanding and supportive. I know I have asked a lot of you while I have been working on this project and I could not be more grateful to you for your aroha and manaakitanga. I love you all so very much. Koia kei a koutou!

To Jacob Edmond, who is my dearest friend and also my husband, thank you for your unwavering support. I could not have completed this thesis without you. Thank you for providing me with a quiet house when I most needed it. Thank you for understanding me. Thank you for having faith in me. I love you.
# Table of contents

Abstract ...........................................................................................................................................i

Acknowledgements ................................................................................................................... ii

Table of contents ....................................................................................................................... iv

List of tables ............................................................................................................................... ix

List of figures ............................................................................................................................... x

Chapter 1. Introduction .............................................................................................................. 1

Origins of the study ................................................................................................................... 1

Focusing in on formative assessment ....................................................................................... 2

Significance of the study ........................................................................................................... 3

Thesis outline ............................................................................................................................. 4

Chapter 2. Educational context and literature review .......................................................... 6

Theoretical underpinning .......................................................................................................... 6

The New Zealand context ......................................................................................................... 7

Tomorrow’s Schools .................................................................................................................. 7

The National Standards ............................................................................................................ 12

Social moderation ..................................................................................................................... 15

Literature review ....................................................................................................................... 17

Dependability, reliability, and validity ..................................................................................... 17

Shared understandings of standards ...................................................................................... 21

The quality of assessment evidence ......................................................................................... 28

The quality of teachers’ judgements ....................................................................................... 30
Assessment for learning ........................................................................................................... 34
Social moderation and teacher learning ............................................................................... 37
Research questions ............................................................................................................... 39
Chapter 3. Methodology ....................................................................................................... 40
The use of case ....................................................................................................................... 40
Selecting and gaining access to schools ............................................................................. 41
Participants ............................................................................................................................ 44
   Central Heights School ..................................................................................................... 46
   Greenville School ............................................................................................................. 46
   Riverside School ............................................................................................................. 47
Data production ..................................................................................................................... 47
   Researcher stance and identity ....................................................................................... 48
   Observation of moderation meetings .............................................................................. 50
   Teacher and principal interviews .................................................................................... 54
   Teacher questionnaire ..................................................................................................... 57
Preparing for analysis: Transcribing observation and interview data ................................ 57
Data analysis .......................................................................................................................... 59
   Initial coding and chronology development .................................................................. 60
   Focused coding ............................................................................................................... 62
Ethical considerations .......................................................................................................... 65
Study and researcher effect .................................................................................................. 66
Chapter conclusion .............................................................................................................. 71
Chapter 4. Moderation as assessment practice: Using involvement in moderation to build teachers’ assessment for learning capability ................................................................. 72

Conceptualising a moderation resource: Developing a system for building students’ assessment capability .............................................................................................................. 73

Constructing assessment resources: Strengthening teachers’ criteria-development skills ................................................................................................................................. 78

Developing a shared perspective: Making connections between assessment and identifying next teaching and learning steps ............................................................................... 83

Chapter conclusion ........................................................................................................... 92

Chapter 5. Moderation and the negotiation of meaning: Conversations about standards, judgement making, and assessment tasks ................................................................. 94

Understandings of standards and assessment criteria ................................................................................................................................. 94

   Syndicate-level understandings ......................................................................................... 95

   School-wide understandings ............................................................................................ 99

Factors affecting the dependability of teacher judgements ................................................. 107

Factors to consider when selecting, designing, and administering assessment tasks ................................................................................................................................. 115

Chapter conclusion ........................................................................................................... 124

Chapter 6. Moderation and assessment identity: Teachers’ accounts of how involvement in moderation shaped their assessment capability ............................................... 126

   Teachers’ views about how involvement in moderation shaped their understandings of writing standards and achievement expectations ...................................................... 126

   Moderation and learning inquiry: Teachers’ views about how moderation shaped their ability to respond to assessment information ................................................. 132

   Identifying what happened as a result of the teaching: Teachers’ views about how moderation shaped their judgement-making ability ............................................. 133
Identifying the implications for future teaching: Teachers’ views about how moderation shaped their ability to identify next teaching and learning steps. 138

Chapter conclusion .................................................................................................................................................. 141

Chapter 7. Moderation as community forming: Using involvement in moderation to create assessment-focused communities of practice .................................................. 143

Characterising a community of practice ................................................................. 143

Using involvement in moderation to learn about assessment ......................... 144

Learning about assessment through involvement in criteria development ....... 144

Learning about assessment through involvement in judgement making .......... 151

Using involvement in moderation to support the formation of assessment-focused communities of practice ................................................................. 154

Limitations of the study ...................................................................................................................... 156

Implications for research................................................................................................. 158

Implications for policy and practice .............................................................................. 159

Recommendation 1: Harness the benefits associated with criteria development ................................................................................................................................. 160

Recommendation 2: Use assessment for learning principles and practices to link moderation with teaching as inquiry and effective pedagogy ........................................... 162

Recommendation 3: Make time for moderation ............................................................ 163

Closing remarks ............................................................................................................................................... 164

References .................................................................................................................................................... 165

Appendix A. Information sheets and consent forms ................................................... 187

Appendix B. Success Criteria developed during moderation meetings at Central Heights School .................................................................................................................. 212

Appendix C. Sample of a Must, Should, Could chart .................................................. 216
<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>Excerpt from field notes</td>
<td>217</td>
</tr>
<tr>
<td>E</td>
<td>Overview of meetings involving teachers in the focus syndicate at</td>
<td>218</td>
</tr>
<tr>
<td></td>
<td>Central Heights School</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>Overview of meetings involving teachers in the focus syndicate at</td>
<td>220</td>
</tr>
<tr>
<td></td>
<td>Greenville School</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Overview of meetings involving teachers in the focus syndicate at</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>Riverside School</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>Interview guide for principals</td>
<td>222</td>
</tr>
<tr>
<td>I</td>
<td>Interview guide for focus syndicate and lead teachers</td>
<td>224</td>
</tr>
<tr>
<td>J</td>
<td>Interview guide for additional lead teacher questions</td>
<td>227</td>
</tr>
<tr>
<td>K</td>
<td>Teacher questionnaire</td>
<td>228</td>
</tr>
<tr>
<td>L</td>
<td>Excerpt from transcription memo</td>
<td>233</td>
</tr>
<tr>
<td>M</td>
<td>Excerpt from a chronology</td>
<td>234</td>
</tr>
<tr>
<td>N</td>
<td>Initial codes</td>
<td>236</td>
</tr>
<tr>
<td>O</td>
<td>Central Heights School moderation plan</td>
<td>245</td>
</tr>
<tr>
<td>P</td>
<td>NEMP assessment task: Torch</td>
<td>246</td>
</tr>
<tr>
<td>Q</td>
<td>Sample of a Tracking Sheet</td>
<td>247</td>
</tr>
<tr>
<td>R</td>
<td>Previously published manuscript</td>
<td>248</td>
</tr>
</tbody>
</table>
List of tables

Table 1. Contextual characteristics of cases in the study........................................44

Table 2. Characteristics of each focus syndicate in the study....................................45

Table 3. Type and amount of data collected for each school (excluding artefacts).....48

Table 4. Summary of data resulting from the observation of moderation meetings....52

Table 5. Summary of teacher and principal interview participation.........................55

Table 6. Number of direct references made to Success Criteria during judgement-
making meetings at Central Heights School.......................................................84

Table 7. The main purposes of moderation identified at each participating school ....91

Table 8. Descriptions of achievement that relate to the use of complex sentences as
specified in the Literacy Learning Progression (Ministry of Education, 2010b) and the New Zealand Curriculum Reading and Writing Standards
(Ministry of Education, 2009c)...............................................................................97

Table 9. Number of judgement-making and criteria-development meetings that the
members of each school’s focus syndicate are known to have participated in
.......................................................................................................................................137

Table 10. Progress by school towards using involvement in moderation to form an
assessment-focused community of practice.........................................................155
List of figures

Figure 1. The number of syndicate and full-staff moderation meetings that teachers participated in each term, by school. CHS = Central Heights School; GS = Greenville School; RS = Riverside School…………………………………….104

Figure 2. The level of agreement that teachers from Central Heights School reported in relation to the statements “Moderation is helping me to clarify my understanding of (a) the National Standards for writing; (b) the Literacy Learning Progressions for writing; and (c) my own school’s writing expectations.”………………………………………………………………………………128

Figure 3. The level of agreement that teachers from Greenville School reported in relation to the statements “Moderation is helping me to clarify my understanding of (a) the National Standards for writing; (b) the Literacy Learning Progressions for writing; and (c) my own school’s writing expectations.”………………………………………………………………………………128

Figure 4. The level of agreement that teachers from Riverside School teachers reported in relation to the statements “Moderation is helping me to clarify my understanding of (a) the National Standards for writing; (b) the Literacy Learning Progressions for writing; and (c) my own school’s writing expectations.”………………………………………………………………………………129

Figure 5. The level of agreement that teachers reported, by school, in relation to the statement “Moderation is making me feel more confident about making dependable judgements of student achievement in writing.” CHS = Central Heights School; GS = Greenville School; RS = Riverside School………..136
Chapter 1. Introduction

In 2010, the New Zealand government introduced a system of achievement standards, called the National Standards, for schools with students in years 1–8 (Ministry of Education, 2009d). This thesis examines the social moderation processes that schools were expected to develop and implement following the introduction of these standards. It investigates whether – and if so, how – New Zealand primary school teachers were able to access professional learning in assessment through their involvement in these social moderation processes.

Origins of the study

In New Zealand, teachers are encouraged to view effective pedagogy as an inquiry process. This process, which sits at the heart of the New Zealand Curriculum (Ministry of Education, 2007), is founded upon the collection and use of sound assessment information. Within this context, teachers gather assessment information using both formal processes, such as the administration of standardised tests, and less formal approaches, including conversing with and observing students. Before instruction, teachers use this assessment information to ascertain what their students already know and what they need to learn next (Ministry of Education, 2011c). During and after teaching, teachers gather and appraise student assessment information, not only to investigate how their teaching has shaped student learning but also to identify next teaching and learning steps. To inquire into their practice in these ways, teachers require the knowledge and skills both to collect sound assessment information and to analyse and interpret this information appropriately.

There are, however, indications that not all New Zealand teachers possess the assessment capability to engage effectively in this important inquiry process. For example, a 2007 study examining the collection and use of assessment information in New Zealand primary schools (Education Review Office, 2007) reported that “the interaction of assessment with teaching and learning” (p. 2) was effective in just over half of the 253 schools in its sample. This is problematic because, despite efforts that the Ministry of Education has made in recent decades to provide schools with access to assessment-focused professional learning initiatives, “a substantial unmet need for assessment professional learning for both teachers and leaders” has been identified (Absolum, Flockton, Hattie, Hipkins, & Reid, 2009, p. 26). The study that is reported on in this thesis examined whether providing teachers with the opportunity to participate in
social moderation, an activity that has been linked with the provision of professional learning in assessment (e.g., Adie, 2010; Hipkins & Robertson, 2012; Reid, 2007; Wyatt-Smith, Klenowski, & Gunn, 2010), might help the sector to meet this need. To do this, it investigated how involvement in moderation shaped the assessment capability of a group of New Zealand primary school teachers. In this country, primary school teachers work with children who are between 5 and 13 years of age, or in years 1–8 of their schooling.

The 2010 introduction of the National Standards assessment regime (Ministry of Education, 2009c), with its associated moderation requirement (Ministry of Education, 2010e), presented an opportune context for this research. Although many New Zealand primary schools had some history of engaging in moderation prior to 2010 (Wylie & Hodgen, 2010), the introduction of the National Standards focused schools’ attention on their moderation processes and increased the prevalence of these activities. Within the National Standards context, schools have been expected to develop and implement moderation processes to improve the dependability not only of teachers’ judgements but also of the assessment information on which these judgements are based (Ministry of Education, 2010f). These moderation processes are also expected to support the principles and practices of assessment for learning (Ministry of Education, 2010e). The study that is reported on here examined how three New Zealand primary schools interpreted these expectations, and identified what teachers learned about assessment from taking part in these moderation activities. The overarching research question that guided this study was: how and what do teachers learn about assessment through their involvement in social moderation?

Focusing in on formative assessment

Student assessment information is gathered for a wide range of reasons. Indeed, Newton (2010) identified 22 purposes for which such information is collected and used. The assessment purposes detailed within his list include informing subsequent teaching, measuring institutional performance, identifying institutional resourcing needs, and evaluating system level performance. Since the introduction of the National Standards (Ministry of Education, 2009c), the New Zealand Ministry of Education (2011c) has explained that information about student performance is required to inform decisions at all levels of the sector. For example, at the school level it is expected that this information will be used to guide both policy and programme reviews and to identify priority areas for professional learning and development. Likewise, assessment
information is required at the sector level to inform a range of decisions, including those related to policymaking and resource allocation.

The interest in this study, though, was in the student assessment information that teachers needed at the classroom and school level in order to inquire into and improve their teaching and learning programmes. The focus, therefore, was on the assessment information that teachers required for formative, rather than summative, purposes. Information gathered both during and at the conclusion of a learning period can be used formatively (Absolum et al., 2009; Gardner, 2010; Ministry of Education, 2011c). As Gardner (2010) explained, “an assessment activity is not inherently formative or summative; it is the purpose and use to which it is put that determines which of these it is” (p. 6). My study focused on whether – and if so, how – involvement in moderation shaped teachers’ ability to collect the assessment information required for formative purposes.

**Significance of the study**

Since the introduction of the National Standards (Ministry of Education, 2009c), New Zealand primary school teachers have spent countless hours engaging in social moderation processes (Ward & Thomas, 2012, 2013, 2015). Within this context, numerous studies have examined the implementation of the National Standards regime (Education Review Office, 2010; Hipkins & Robertson, 2012; Thomas & Ward, 2011; Thrupp, 2013; Thrupp & Easter, 2012; Thrupp & White, 2013; Ward & Thomas, 2012, 2013, 2015; Wylie & Bonne, 2014; Wylie & Hodgen, 2010). Of these, only one study (Hipkins & Robertson, 2012) has explored what teachers learned from their involvement in the moderation processes that are associated with this assessment system. Building on the work of Hipkins and Robertson, which was based on a single case study, the research project that is reported on in this thesis involved the teachers at three New Zealand primary schools. As well as extending our understandings of what teachers can learn about assessment through their involvement in moderation, this multi-case study has produced insights into the conditions and contexts that enable teachers to use their participation in moderation to establish assessment-focused communities of practice. Because of this, these findings will be of particular interest to schools that are using their involvement in the Ministry of Education’s new Communities of Learning initiative (Ministry of Education, 2016a) to revitalise and extend their moderation processes. More generally, these findings should enable New Zealand primary school teachers, and indeed teachers working in other contexts, to
utilise the time that they spend taking part in moderation activities to strengthen their assessment capability.

**Thesis outline**

This first chapter has served to provide a broad-brushstrokes introduction to the study that is reported on in this thesis. Further details about the context within which this study occurred are presented in Chapter 2. As well as providing an overview of New Zealand’s educational landscape, this chapter introduces the concept of social moderation and explains how this assessment activity was specified for National Standards purposes. After defining the terms *dependability, reliability,* and *validity,* a review of the literature on social moderation is presented and the study’s research questions are stated. Parts of Chapter 2 were used as the basis for an article (Smaill, 2013) that was published in the journal *Assessment in Education: Principles Policy & Practice.* As the sole author of this article, which is entitled “Moderating New Zealand’s National Standards: Teacher learning and assessment outcomes,” I have been granted permission to include it at the end of this thesis (see Appendix R).

Chapter 3 explains how both social constructionism (Crotty, 1998) and Wenger’s (1998) social theory of learning informed the methods that were employed within this multi-case study. This chapter describes how each of these theories guided the ways in which data were both produced and analysed. As well as introducing the processes that were used when selecting the three participating schools, this chapter describes the methods that were used to produce data and provides a rationale for the inclusion of observations, interviews, and a questionnaire. It then explains how key elements of grounded theory (Charmaz, 2006) informed the analytic approach that was used. The chapter concludes by outlining the ethical considerations that were associated with this project, including the ways in which this study might have shaped the social moderation activities that it set out to examine.

Three findings chapters follow. Chapter 4 draws on Wenger’s (1998) conception of practice to examine how teachers used their involvement in social moderation to deepen their understandings of assessment for learning principles and practices. It identifies how taking part in the development of moderation resources and processes enabled teachers to garner these learning opportunities. Likewise, it highlights the ways in which schools utilised their existing assessment expertise to maximise the professional learning benefits that participating in moderation afforded teachers. Chapter 5 explores how and what teachers’ experiences of moderation taught them
about factors that can affect the dependability of assessment information. Here, Wenger’s notion of meaning is used to gain insights into how involvement in moderation shaped this aspect of teachers’ assessment capability. This chapter also investigates how the leadership of each school’s moderation processes affected teacher learning. Drawing on teachers’ accounts of how and what they learned about assessment through participating in moderation, Chapter 6 utilises Wenger’s conception of identity to explore how involvement in moderation shaped teachers’ perceptions of their assessment capability. This chapter draws attention to the reports that teachers provided about the ways in which developing and using school-specific assessment criteria improved their assessment capability.

The findings from this study are woven together in Chapter 7. Here, Wenger’s (1998) conception of community is used to explain why those teachers who used their involvement in moderation to form a school-wide community of practice seemed to have the greatest opportunity to learn about assessment from these experiences. This final chapter examines the ways in which this study’s findings augment existing understandings about how and what teachers learn about assessment through their involvement in social moderation. As well as acknowledging the study’s limitations, this chapter also explores its implications for research, policy, and practice.
Chapter 2. Educational context and literature review

A key goal of this study was to examine if and how teachers might access professional learning in assessment through their involvement in the school-based social moderation processes that are associated with New Zealand’s National Standards assessment regime. This chapter provides both the context and the justification for this study. After introducing the social constructionist underpinning (Crotty, 1998) that informs this study, it provides a brief overview of New Zealand’s educational assessment context. Here, particular attention is paid to the role that assessment for learning has come to play within this landscape. Next, the concept of social moderation (Linn, 1993; Maxwell, 2010) is introduced and details are provided about how this assessment activity has been specified for the purposes of National Standards. Because these moderation processes are intended to improve the quality of the assessment information that teachers generate (Ministry of Education, 2010f), the terms dependability, reliability, and validity are then defined, and their function within this thesis is explained. The rest of the chapter reviews the existing literature on social moderation. In particular, it identifies what is already known about the ways in which involvement in these processes informs two aspects of teachers’ assessment capability. These are teachers’ understandings of factors that can affect the dependability of assessment information and teachers’ understandings of the principles and practices of assessment for learning. As well as allowing some conclusions to be drawn about how taking part in social moderation might shape these aspects of teachers’ assessment capability, this also serves to highlight the gaps in the existing research. The chapter concludes by presenting the research questions that this study seeks to answer.

Theoretical underpinning

As explained in the introductory chapter, the overarching research question that guides this study is: how and what do teachers learn about assessment through their involvement in social moderation? Embedded in this question are a series of assumptions about the nature of reality (ontology), knowledge (epistemology), and learning. These assumptions reflect the social constructionist perspective that underpins this study. From this perspective, all individuals construct knowledge, or meaning, as they interact with and interpret the world (Burr, 1995; Crotty, 1998). These interactions, which inevitably render multiple and varied interpretations (Creswell, 2013), are shaped by both culture and history. As Burr (1995) has explained, “the ways in which we
commonly understand the world, the categories and concepts we use, are historically and culturally specific” (p. 3). Therefore, any attempt to explore how and what teachers learn about assessment through their involvement in moderation must take into account the historical and cultural contexts within which this activity occurs. This is addressed in the next section, which provides a brief summary of important events in New Zealand’s recent educational history.

The New Zealand context

Although the National Standards assessment landscape provides the immediate backdrop to this study, the provision of a broader historical perspective affords useful insights into the learning that teachers experience through their involvement in moderation. Beginning with the introduction of the Tomorrow’s Schools reforms in the late 1980s (Department of Education, 1988), which some of this study’s participants \( n = 8 \) would have experienced firsthand, the following section provides an overview of significant events in New Zealand’s educational history. Spanning the years from 1988 to 2010, this overview provides contextual information about the administration and organisation of New Zealand’s primary schools and describes factors that are likely to have shaped teachers’ assessment practices during these years. In particular, it charts the emergence and evolution of “assessment for learning” within the New Zealand primary school sector (Ministry of Education, 2011c, p. 9). Additionally, as this thesis examines how involvement in the moderation of student writing shaped teachers’ assessment capability, this section also introduces those centrally developed resources associated with the teaching and assessment of writing that were implemented during this period.

Tomorrow’s Schools

The 1988 introduction of Tomorrow’s Schools initiated sweeping changes to the administration of education within New Zealand’s schooling sector (Department of Education, 1988). These changes had implications for the ways in which New Zealand primary school teachers were expected to assess and report on student achievement. As a result of these reforms, a policy-focused Ministry of Education was created and the responsibility for administering and managing individual schools was removed from regional education boards and devolved to each school’s local community (Ministry of Education, 2010h). To this day, a board of trustees comprised of the principal, a teacher representative, and a number of parent-elected board members is responsible for the governance of each of New Zealand’s state and state-integrated primary schools.
As well as promoting increased parental and community involvement, the Tomorrow’s Schools reforms were, according to the Prime Minister of the time, expected to “lead to improved learning opportunities for the children of this country” (Department of Education, 1988, p. iv). Within this new educational landscape, each school’s board of trustees assumed responsibility for ensuring that these improvements occurred (Department of Education, 1988).

Around the time that Tomorrow’s Schools was introduced, there was, in the words of Gipps (1994), “an explosion of developments in assessment” (p. 1). Many of these developments were reflected in the assessment policies that were introduced to New Zealand primary schools during the late 1980s and throughout the 1990s. The move in the primary sector from a norm-referenced assessment model towards a standards-based system provides a case in point (Ministry of Education, 1990). Until 1989, teachers had rated student achievement in each subject on a 5-point scale based on the normal curve. Recognition that this approach did not readily enable teachers to identify future learning needs motivated the move towards a standards-based system (Ministry of Education, 1990).

The notion of standards-based, or as Sadler (1987) first coined it, “standards-referenced” (p. 193), assessment had been developed just a few years earlier in response to increased dissatisfaction with the limitations of norm-referenced assessment models. Sadler proposed standards-referenced assessment as an approach that would support the development of students’ evaluative expertise and enable teachers to generate the assessment information that they required for both improving learning and summative reporting. The idea that assessment information should be used for improving learning was consistent with the Ministry of Education’s (1990) “principles of assessment for better learning” (p. 7). These included the statement:

The primary purpose of assessment should be to provide information which can be used to identify strengths and guide improvement. In other words, it should suggest actions which may be taken to improve the educational development of students and the quality of education programmes. (p. 8)

The publication of these principles signalled the emergence of the Ministry of Education’s commitment to assessment for learning (Ministry of Education, 2011c). In New Zealand, the phrases formative assessment and assessment for learning have come to be used interchangeably (Ministry of Education, n.d.-f).
It was within this improvement-oriented assessment landscape that New Zealand primary school teachers were expected to begin using seven new progressively introduced National Curriculum Statements (henceforth, curriculum statements; Ministry of Education, 1992, 1993a, 1994b, 1995, 1997, 1999a, 2000). The content of each of these curriculum statements related to a particular learning area (for example science, mathematics, and English) and described the achievement that students were expected to attain in that area during years 1–13 of their schooling. Rather than specifying achievement in terms of school year levels, these statements used a system of curriculum levels for this purpose. Each of these statements provided a list of broad achievement objectives that explained the 8 levels of attainment that students were expected to achieve (Ministry of Education, 1994a). The expectation was that teachers would break down these broad achievement objectives into more specific learning outcomes. After teaching and assessing children in relation to these more specific outcomes, teachers were required to make qualitative judgements about whether each student had met the broad objectives.

In 1993, a revised version of the National Education Guidelines came into effect (Ministry of Education, 1993c). These guidelines, which set out the accountability requirements for schools, specified how teachers were to use the curriculum statements to inform their teaching and assessment of students. They also clarified the role of boards of trustees in relation to the learning and achievement of students. Because these guidelines were broad and not particularly prescriptive, schools were afforded relative freedom in deciding how to implement (Ministry of Education, 1998), assess, and report against the new curriculum statements (Absolum et al., 2009).

In 1995, to help schools improve their assessment and reporting practices within this new and rapidly evolving educational environment, the Ministry of Education began funding the Assessment for Better Learning (ABeL) professional development programme (Peddie, 2000). During the 7 years that the ABeL programme was offered (1995–2001), about 400 schools took part per year. Eppel (1998) explained that the ABeL programme was intended to help teachers and principals develop school-based assessment practices that reflected both the policy directions in the New Zealand Curriculum Framework (Ministry of Education, 1993b) and the requirements specified in the National Education Guidelines (Ministry of Education, 1993c). Although the specific goals of this programme are unclear, one review (Peddie, 2000) indicated that the ABeL programme provided participants with some opportunities to develop their understandings of both formative assessment and the nature of effective feedback. This
suggests that efforts were made to integrate burgeoning understandings about the qualities of effective feedback (Crooks, 1988) and the role and function of formative assessment (Black & Wiliam, 1998; Gipps, 1994; Sadler, 1989) into the ABeL programme. Although the Peddie (2000) report recommended that this programme needed to strengthen its delivery in these areas, the ABeL programme provides an example of the Ministry of Education’s flourishing commitment to the principles of assessment for learning.

In the late 1990s, the Ministry of Education became concerned that primary school teachers lacked the assessment tools required to generate information about how their students were performing in relation to national expectations (Ministry of Education, 1998, 1999b). In response to this, a variety of new assessment resources were developed. These included, but were not limited to, the New Zealand Curriculum Exemplars (henceforth, the curriculum exemplars; Ministry of Education, 2003a, 2003b, 2003c, 2004a, 2004b, 2004c, 2004d) and a resource called the Assessment Tools for Teaching and Learning (asTTle; Hattie, Brown, Keegan, & Mackay, 2004). The curriculum exemplars, which were developed in conjunction with teachers, provided authentic, annotated samples of student work that represented each of the levels specified in the seven national curriculum statements (Poskitt, Brown, & Taylor, 2003). For written language, these exemplars also included a series of matrices that set out progress indicators (Ministry of Education, 2003b). These indicators illustrated the English curriculum achievement objectives for levels 1–5. Thus, they exemplified the range of achievement that was expected in written language during years 1–8.

The development of the curriculum exemplars was important because the achievement objectives specified in New Zealand’s curriculum statements are an example of what Sadler (1987) has described as “verbal descriptions” (p. 201) of standards. Because such standards are specified in linguistic terms, their meanings are “always to some degree vague or fuzzy” (p. 202). Given this issue, the provision of exemplars, or concrete reference points, can help teachers and students to clarify their understandings of the expected standards (Sadler, 1987, 1989). This was an intended function of the New Zealand curriculum exemplars, which it was hoped would “illustrate key features of learning, achievement, and quality at different stages of student development; help students and teachers to identify the next learning steps; and guide teachers in their interpretation of curriculum levels” (Ministry of Education, 2003b, p. 1). These goals suggest that a desire to strengthen the skills and knowledge
associated with the principles of assessment for learning may have motivated the
development of the curriculum exemplars.

In contrast, a somewhat different goal appears to have catalysed the
development of the asTTle resource (Hattie et al., 2004). The origins of asTTle can be
traced back to a 1998 green paper that identified the need for mandatory, national,
externally referenced standardised tests (Ministry of Education, 1998). With regard to
these tests, this green paper explained that “the key purpose of the proposed tests is to
provide schools with information that enables them to compare their students’
achievement with national levels of achievement and with those of similar students
nationwide” (p. 26). Although these tests were presented as a mechanism for providing
teachers and principals with the information that they required to improve student
learning opportunities (Ministry of Education, 1998), some questioned the likelihood
that the data these tests generated would be used for this purpose (Hattie, 1998). Indeed,
the proposal to introduce these national, externally referenced tests met with very strong
opposition (Gilmore, 1998).

Notably, schools were initially provided with access to the resulting asTTle
resource (Hattie et al., 2004) on a voluntary, rather than a mandatory, basis (Brown,
2013). More recently, however, schools have been required to use its successor, named
e-asTTle (Ministry of Education, n.d.-e), to participate in certain Ministry of Education-
funded initiatives (Consortium for Professional Learning, 2016). The asTTle resource
(Hattie et al., 2004), which was first made available to all schools in 2005, provided
teachers with an electronic tool for developing paper-and-pencil tests (Brown, 2013).
These tests, which could be created to assess reading, writing, and mathematics,
provided information about a student’s achievement in relation to levels 2–6 of the
relevant curriculum statement. In writing, teachers generated this information by using a
series of rubrics (e.g., Ministry of Education, 2012) to evaluate their students’ responses
to a common writing assessment. In keeping with the original intent of this initiative,
teachers could input the resulting data into the asTTle software and generate
information that allowed them to analyse student achievement in relation to both
curriculum levels and population norms (Ministry of Education, n.d.-c). To complement
this, asTTle also provided a link to an online catalogue of classroom resources, called
“What Next” (Ministry of Education, n.d.-j). This resource, which has continued to be
updated, enables teachers to use student assessment information to identify next
learning steps.
A variety of Ministry-funded professional development opportunities were provided to schools to support the effective use of the asTTle system. As well as having the option to attend specific asTTle workshops (Evaluation Associates Ltd., 2007), schools could also strengthen their use of this resource through participating in the Assess to Learn (AToL) project (Poskitt & Taylor, 2008). Building understanding of asTTle, however, was just one of the intended functions of the AToL initiative. This initiative, which succeeded the ABeL programme in 2002, was introduced to provide teachers and principals with in-depth professional learning in the use of assessment for learning principles and practices. To achieve this goal, the AToL project provided participants with the opportunity to develop their skills in areas such as providing feedback and feed-forward in the form of next steps, crafting learning intentions and success criteria, fostering student self-assessment, using assessment information to adjust teaching and learning programmes, making effective use of assessment tools such as the curriculum exemplars and asTTle, and utilising student work as a basis for discussions. Moreover, other professional learning initiatives that were offered around this time, such as the Literacy Professional Development Project, also supported the development of many of these skills (McDowall, Cameron, Dingle, & Gilmore, 2007).

During the 8-year period in which the AToL project was offered to schools, New Zealand’s seven curriculum statements (Ministry of Education, 1992, 1993a, 1994b, 1995, 1997, 1999a, 2000) were replaced by a single-volume curriculum (Ministry of Education, 2007). Sewell (2007) has explained that the revised New Zealand Curriculum (henceforth, the curriculum) underwent an extensive development period during which it benefited from the involvement of “more than 15,000 students, families, schools, communities and educators” (p. 14). It has been noted (Crooks, 2011; Hipkins & Robertson, 2011) that quite a different process was employed during the development of those documents associated with the national assessment system that was introduced shortly after this curriculum.

The National Standards

In 2010, the New Zealand government brought in a system of achievement standards called the National Standards, for English-medium schools with students in years 1–8 (Ministry of Education, 2009d). The following year, equivalent standards were introduced for Māori-medium primary schools (Ministry of Education, 2010g). These standards (Ministry of Education, 2009b, 2009c) were developed rapidly (Hipkins & Robertson, 2011) and few teachers were provided with the opportunity to be involved in
this process (Crooks, 2011). New Zealand’s National Standards (Ministry of Education, 2009b, 2009c), which were introduced to enable students to meet the demands of the curriculum (Ministry of Education, 2007), specify yearly achievement expectations in reading, writing, and mathematics. They consist of descriptors and annotated exemplars of student work that are linked to school year levels. As such, they are a form of what Maxwell (2002a) has described as standards used as “quality benchmarks” (p. 2). During the assessment of writing, teachers are also expected to make use of a companion document (Ministry of Education, 2009c), entitled the Literacy Learning Progressions (Ministry of Education, 2010b), to inform their judgements of student achievement. Notably, the standards that are specified within this companion document are also in the form of quality benchmarks.

Mirroring a global trend in education (Klenowski & Wyatt-Smith, 2010a), New Zealand’s National Standards are expected not only to improve educational outcomes (Ministry of Education, 2009c) but also to strengthen accountability (New Zealand National Party, 2011). Although teachers are required to assess all students against this set of common standards, the approach to gathering assessment information is not tightly specified, and there is no national test (Ministry of Education, 2011c). In fact, because the use of common assessment tasks might not provide all students with the opportunity to perform optimally (Ministry of Education, 2011d), teachers have been encouraged to tailor their selection of both self-generated and centrally developed assessments to the needs of individual students (Ministry of Education, 2010f). Teachers then use each student’s unique evidence base to inform a qualitative judgement of that student’s achievement in relation to the National Standards (Ministry of Education, 2010f). These summative judgements, called “overall teacher judgements” (OTJs), must be made twice yearly for each student in relation to the standards specified for reading, writing, and mathematics (Ministry of Education, 2010a). Although New Zealand primary school teachers have a long history of gathering and using student assessment information to inform their summative judgements of student achievement (Department of Education, 1989; Ministry of Education, 1994a), the introduction of the National Standards significantly altered the context within which these judgements were reported.

Prior to 2010, and since the introduction of self-management in 1988, New Zealand’s primary schools had considerable freedom to choose how to approach assessment and reporting with respect to the National Education Guidelines (Absolum et al., 2009; Ministry of Education, 2010h). Amendments made to the National
Administration Guidelines since the introduction of the National Standards, however, resulted in changes that signalled a major shift in New Zealand’s assessment and reporting policy. Schools must now report annually on “the numbers and proportions of students achieving at, above, below, or well below the standards” (Ministry of Education, 2009d, p. 2). Moreover, since 2012 each school’s National Standards data have been made publicly available on a New Zealand government website (Thrupp, 2013; Thrupp & White, 2013). It is in this environment of increased accountability that the Ministry of Education (2010e) has expected teachers to participate in intra-school, and “where appropriate” (p. 2) inter-school, moderation processes as part of system-level efforts to improve the dependability of student assessment information.

The need to moderate arises in part because of the nature and quality of achievement standards. New Zealand’s National Standards are a form of what Sadler (1987) has described as verbal descriptions or qualitative rubrics. As previously explained, such standards always possess an inherent “fuzziness” (Sadler, 1987, p. 202), and it is widely recognised that they require interpretation (Connolly, Klenowski, & Wyatt-Smith, 2012; Klenowski & Adie, 2009; Klenowski & Wyatt-Smith, 2010a, 2010b; Maxwell, 2002b, 2002a, 2009, 2010; O’Donovan, Price, & Rust, 2004; Sadler, 1987, 2009; Wyatt-Smith et al., 2010). Given that an assessor’s interpretation of a standard is shaped by his or her unique evaluative experience (Sadler, 1987), and indeed the assessment context (Sadler, 2009), these interpretations inevitably differ. Because of this, a number of conditions must be fulfilled to help assessors clarify their interpretations of standards and develop common understandings of their meanings (Klenowski & Wyatt-Smith, 2010a; O’Donovan et al., 2004; Sadler, 1987). Assessors first require exemplars or concrete examples demonstrating the specified levels of achievement. Although necessary, exemplars are not on their own sufficient to ensure that assessors develop shared understandings of achievement standards. Crucially, assessors also need opportunities to participate in moderation processes. The learning opportunities that involvement in these processes afforded New Zealand primary school teachers are the focus of this thesis.

Moderation is established in New Zealand’s secondary schools (Hipkins, 2010; Hipkins & Robertson, 2011), where the outcomes of standards-based National Certificate of Educational Achievement (NCEA) assessments influence students’ tertiary and employment opportunities (Ministry of Education, 2011c). Because of the high stakes associated with these assessments and the importance of ensuring their integrity, NCEA moderation involves both internal and external processes (Ministry of
These include the submission of assessment materials for pre-moderation, school-based moderation or co-marking of student work, and the submission of marked assessments for external moderation.

**Social moderation**

Given that New Zealand’s National Standards are associated with relatively lower stakes than NCEA assessments, a somewhat less prescribed moderation system has been adopted at the primary level (Ministry of Education, 2011c). Instead, New Zealand’s primary school teachers are required to engage in a process that is most commonly described as social moderation (Linn, 1993; Maxwell, 2010). This approach, which has also been referred to as group moderation (Gipps, 1994; Harlen, 1994), involves the comparison and alignment of participants’ judgements of student work in relation to stated standards (Connolly et al., 2012; Klenowski & Wyatt-Smith, 2014; Maxwell, 2010). During this participative process, assessors clarify their interpretations of qualitative descriptors and develop shared understandings of their meaning (Harlen, 1994; Klenowski & Wyatt-Smith, 2014; Maxwell, 2002b, 2010). Maxwell (2010) emphasised that social moderation should “be framed as a normal and positive consultative process that values consensus and builds confidence” (p. 469). Consensus is achieved when assessors reach agreement about the standard awarded to a student’s assessment (Connolly et al., 2012; Klenowski & Wyatt-Smith, 2014; Maxwell, 2002b, 2010). Linn (1993) posited that widening social moderation processes to an inter-school level can both broaden consensus and strengthen public confidence, and New Zealand primary schools have been encouraged to establish inter-school moderation networks (Ministry of Education, 2010e). Henceforth, the terms *moderation* and *social moderation* are used interchangeably.

Social moderation can serve a dual function (Harlen, 1994; Klenowski & Wyatt-Smith, 2014; Maxwell, 2002b, 2010), fulfilling what Harlen (1994) has described as both quality control and quality assurance purposes. Maxwell (2010), who used the terms “moderation for accountability” and “moderation for improvement” (p. 457) to make a similar distinction, also suggested that the goal is for each of these functions to complement the other. As the terminology implies, quality control or accountability functions (which will henceforth be referred to as accountability functions), are product oriented and involve the adjustment or verification of assessment outcomes to ensure fairness and comparability (Harlen, 1994; Maxwell, 2002b, 2010).
In contrast, improvement or quality assurance functions (which will from now on be referred to as improvement functions) are largely process oriented. Moderation for improvement strengthens the capability of participants to conduct appropriate assessments and arrive at consistent and comparable judgements (Harlen, 1994; Maxwell, 2002b, 2010). Klenowski and Wyatt-Smith (2014) argued that this form of moderation should “improve teachers’ assessment and pedagogic practice” (p. 74), and thereby contribute to lifting student outcomes. Viewed in this way, the principles of moderation for improvement are integral to effective pedagogy. Accordingly, moderation for improvement is consistent with the notion of “teaching as inquiry” that is described in the New Zealand Curriculum (Ministry of Education, 2007, p. 35). In terms of functionality, it is moderation for improvement that is of primary interest within this thesis.

The information that the Ministry of Education has provided about moderation within the National Standards context indicates that these processes are expected to fulfil both accountability and improvement functions. Evidence that these procedures are expected to serve an improvement function can be found in the Ministry of Education’s (2011b) suggestion that teachers should view National Standards moderation as a “learning exercise” (p. 38) rather than just an opportunity to check the accuracy of their judgements. Consistent with this orientation, the Ministry of Education has encouraged teachers to initiate their moderation work at the beginning of the teaching and learning cycle, as they plan for teaching and assessment. Moreover, it has also emphasised the existence of links between moderation processes and the principles of assessment for learning (Ministry of Education, 2010e).

At the same time, there are clear indications that the Ministry of Education also expects these moderation processes to serve an accountability function. Specifically, schools have been advised that a lack of improvement in the percentage of agreement between teachers’ judgements is a signal that moderation processes are not achieving the “desired outcome” (Evaluation Associates, 2010, p. 1). Linn (1993) suggested that public confidence in such moderation systems is typically achieved through provision of information about the degree of inter-rater agreement. Mirroring this sentiment, the Ministry of Education (2011d) explained that National Standards moderation processes should provide “assurance to parents and others that interpretations of students’ achievements are in line with other professionals [sic]” (p. 12).

In contrast with New Zealand’s NCEA system, the National Standards system has no mechanisms in place for externally moderating or verifying either teacher-
generated assessment tasks or teacher judgements. Teachers sometimes view external moderation negatively because it can serve an accountability function (Hipkins, 2010). Given the right conditions, however, external moderation can also provide teachers with valuable feedback about their understanding of standards and the quality of work associated with different levels of achievement (Hipkins, 2010; Maxwell, 2010). Within the National Standards context, the lack of such an external moderation process eliminates a feedback loop that might otherwise have both facilitated and informed teacher professional learning. Although it is acknowledged that a relatively new online resource, called the Progress and Consistency Tool (PaCT; Ministry of Education, 2015), has the potential to provide teachers with some feedback about their judgement making, this tool was only launched in late 2014 (Ministry of Education, 2014). It was therefore not available to schools at the time of data collection for this study.

**Literature review**

The moderation processes that are associated with New Zealand’s National Standards are intended to improve the dependability of both teachers’ judgements and the assessment evidence on which these judgements are based (Ministry of Education, 2010f). During their involvement in these processes, teachers are also expected to engage with the principles and practices of assessment for learning (Ministry of Education, 2010e). This thesis investigates whether taking part in moderation might strengthen each of these aspects of teachers’ assessment capability. Specifically, it examines whether – and if so, how – taking part in the moderation processes that are associated with the National Standards could enable teachers to learn about factors that affect the dependability of assessment information. It also explores whether – and if so, how – involvement in these moderation processes could strengthen teachers’ understandings of the principles and practices of assessment for learning. The upcoming review of the literature examines what is already known about how involvement in moderation shapes each of these aspects of teachers’ assessment capability. Throughout this review, attention is drawn to areas in which further research is needed. Given that the terms dependability, reliability, and validity feature prominently, a brief explanation of each of these concepts is now presented.

*Dependability, reliability, and validity*

The notion of dependability, which can be expressed as the sum of reliability and validity (James, 1998), is often used to describe the quality of assessment information in
qualitative research. This is true of its application within the National Standards context, where it has been specified that a dependable assessment is one that has “both high validity and reliability” (Ministry of Education, n.d.-c, Dependability, para. 1). Harlen (2005) explained that the concept of dependability was developed to acknowledge the interaction that exists between reliability and validity. New Zealand’s Ministry of Education referred to this interaction as sometimes involving “a trade-off between validity and reliability” and acknowledged that “formal assessments may be more reliable, while informal assessments may be more valid” (Ministry of Education, n.d.-c, Dependability, para. 1). Others have recognised the role that the purpose of an assessment should play in determining the relative importance of reliability and validity. For example, Harlen (2005) suggested that where teachers’ assessments are used for summative purposes in place of testing and have a goal of protecting construct validity, it would be important to consider “what is the highest optimum reliability that can be reached whilst preserving construct validity” (p. 248). Within this thesis, a simpler definition of dependability is employed. In keeping with the New Zealand Ministry of Education’s (n.d.-g) usage, the term is used purely to acknowledge both reliability and validity in combination. Given the importance of each of these concepts, these terms are now defined and the factors that are known to affect both reliability and validity are explored.

The concept of reliability is used to refer to the consistency of the results of an assessment (Darr, 2005a; Koretz, 2009). The reliability of both an assessment and those scoring that assessment can be measured. With regard to scoring or judgement making, the term reliability is typically used to describe the level of agreement that exists between markers. As previously explained, achieving inter-rater agreement or consensus is often a goal of moderation activity (Connolly et al., 2012; Klenowski & Wyatt-Smith, 2014; Maxwell, 2002b, 2010). It has been recognised that inter-rater reliability is affected by the degree to which assessors are able to develop shared understandings of stated standards and assessment criteria, and it is understood that involvement in moderation can assist with this (Klenowski & Wyatt-Smith, 2010a, 2014; Maxwell, 2010). Therefore, the development of shared understandings of standards and criteria emerges as a factor that can affect the reliability of teacher judgements.

For the reliability of an assessment itself to be measured, a student must respond to more than one question, item, or task (Parkes, 2012). Without such replication, it is not possible to estimate the consistency of a student’s response or the extent to which a
generalisation can be reached on the basis of the student’s performance. For this reason, increasing the number of items or tasks on an assessment, or having students undertake similar assessment tasks on more than one occasion, is generally understood to increase the reliability of an assessment (Crooks, Kane, & Cohen, 1996; Darr, 2005a; Parkes, 2012; J. Smith, 2003). Inconsistencies in the ways in which an assessment is administered or presented to students can also affect reliability. Variation in things such as the provision of task instructions or the time allowed for an assessment to be completed can cause variability in student scores (Crooks et al., 1996; Darr, 2005a; James, 1998). Therefore, when considering factors that affect the reliability of assessment evidence, both the administration (or conditions of assessment) and the sufficiency of the information gathered deserve attention.

The validity of an assessment is a measure of the quality of the inferences and decisions that can be made on the basis of that assessment (Kane, 2006; Messick, 1989). Although validity is not an inherent characteristic of a test or a task itself (Koretz, 2009), factors that are known to affect validity require careful consideration during the selection and construction of assessment tasks (Darr, 2005b). Harlen (2005) noted that this is because “validity refers to what is assessed and how well this corresponds with the behaviour or construct that it is intended to test or assess” (p. 247). Therefore, when either developing an assessment task or considering the suitability of an existing task for a given purpose, it is necessary to think about the degree to which that task measures what it is intended to measure.

Consideration of the construct, or the specific psychological characteristic or trait that an assessment is meant to measure, should come to the fore during the design or selection of an assessment task. With regard to classroom assessment, Brookhart (2003) emphasised the importance of “defining the construct in its instructional and contextual sense” (p. 10). Achieving this requires a robust understanding of the construct in question and how that construct might be exhibited (Darr, 2005b). Because definitions of constructs are typically conveyed to teachers through documents that specify the achievement standards for a given learning area, the depth of understanding that teachers possess in relation to those standards is a factor that can affect the design and selection of assessment tasks. In turn, this has implications for the inferences, or judgements, that are arrived at on the basis of the information that these tasks generate. Examining the content of an assessment as part of a validity argument involves determining how well that assessment represents or samples the intended learning domain (Darr, 2005b). In this regard, Stiggins (2008) stated that “assessments that
appropriately cover the material to be learned are said to meet standards of content validity” (p. 41). An assessment task might fail to assess parts of the target domain if the developer has not identified these parts of the domain and therefore has not covered these areas in the assessment (Crooks et al., 1996). Because of this, the degree of alignment that exists between the stated learning intentions and the skills, knowledge, and abilities that an assessment elicits is a factor that affects validity (Darr, 2005b).

Finally, the opportunity that students are afforded to demonstrate their knowledge and ability in the specified area is a key feature of validity (Harlen, 1994). A variety of factors can affect whether an assessment provides students with the opportunity to perform optimally. For example, Harlen (1994) explained that both the meaning that children perceive for a task and their interest in it can influence performance. Similarly, Crooks et al. (1996) have specified that students may not put maximum effort into an assessment if they perceive that the task is artificial or lacks personal relevance. Likewise, Wise and Smith (2011) have developed a demands-capacity model of test-taking effort that addresses not just tests, but the items on those tests. They have explained that having limited or negative experiences of the context underpinning a test item could have a detrimental effect on the effort expended by a student.

To appraise the validity of a judgement, the quality of the inferences and interpretations on which that judgement was based must be evaluated. A number of factors can threaten the validity of an assessor’s interpretation of assessment evidence. For example, inappropriate judgements can be arrived at if an assessor does not fully understand either the assessment information or its limitations (Crooks et al., 1996). Likewise, the intrusion of irrelevant contextual information (Harlen, 1994), which may include teacher biases (Crooks et al., 1996), can also threaten the validity of a judgement. The upcoming review of the literature examines the extent to which involvement in moderation draws teachers’ attention to the effect that factors like these can have on the dependability of assessment information. Because it has been shown that teachers’ understandings of stated standards shape both their judgements of student achievement and their selection and development of assessment tasks, this review of the literature begins with an evaluation of the opportunities that moderation processes can afford teachers for establishing shared understandings of standards.
Shared understandings of standards

Numerous studies have reported on how New Zealand primary schools have responded to the introduction of the National Standards (Education Review Office, 2010; Hipkins & Robertson, 2012; Thomas & Ward, 2011; Thrupp, 2013; Thrupp & Easter, 2012; Thrupp & White, 2013; Ward & Thomas, 2012, 2013, 2015; Wylie & Bonne, 2014; Wylie & Hodgen, 2010). Each of these studies provides some details about the moderation processes that schools are employing within this context. For example, the Education Review Office report (2010) noted that these moderation processes seemed to be working the most effectively at schools in which teachers had previously engaged in “regular professional learning conversations about assessment data” (p. 10). Yet none of this research specifically examined how teachers’ involvement in moderation shaped their development of shared understandings of stated standards. It is, however, possible to draw some inferences about this from other findings that these studies have reported. Further insights can also be gleaned from those studies that examine social moderation within different geographic contexts (Adie, 2010; Black, Harrison, Hodgen, Marshall, & Serret, 2011; Estyn, 2010; Klenowski & Adie, 2009; Klenowski & Wyatt-Smith, 2010b; Reid, 2007; Watty et al., 2014; Wyatt-Smith et al., 2010).

Since the introduction of the National Standards (Ministry of Education, 2009c), the Ministry of Education (2011a) has encouraged teachers to use their involvement in moderation to draw upon resources, such as the Reading and Writing Standards (Ministry of Education, 2009c) and the Literacy Learning Progressions (Ministry of Education, 2010b), to develop local assessment criteria. Very little is known about how engaging in such processes shapes the development of shared understandings of standards and criteria. Shortly after the introduction of the National Standards, one study (Thomas & Ward, 2011) identified the resources that teachers had used to create student performance criteria for moderation. More recently, it has been noted that many New Zealand primary teachers have continued to use these school-developed resources to inform their moderation conversations (Ward & Thomas, 2015, 2016). These researchers did not, however, explore whether the experience of developing or using

---

1 In late 2014 (A. Carlisle, personal communication, September 11, 2015), the Ministry of Education added some additional advice about the use of school-defined standards to its own Assessment Online website (Ministry of Education, n.d.-d). This advice stated that the notion that “judgements should be made in relation to school-defined standards” was a “misunderstanding” (Misunderstanding 3). Despite this, their earlier advice about using involvement in moderation to develop local success criteria is still available online (Ministry of Education, 2011a, 2011b).
such local criteria enabled teachers to arrive at shared understandings of these descriptions of achievement.

Only one other study of moderation (Reid, 2007) has documented the processes that teachers employed as they collaborated to develop assessment criteria. This study described some of the issues that Scottish teachers encountered as they used a set of National Assessment Bank criteria to create a detailed rubric for imaginative story writing. The issues that were noted here included teachers tending to focus their discussions on those criteria that described more readily measurable features of writing, such as the accuracy of spelling and punctuation, or the degree of variation in sentence structure. Although Reid noted that involvement in the criteria-development process strengthened the teachers’ understandings of the writing progression, this research did not provide any insights into the interpretive acts that the Scottish teachers engaged in as they used the National Assessment Bank criteria to create more detailed local criteria.

One other study (Hargreaves, Galton, & Robinson, 1996) examined how taking part in an activity that was very similar to criteria development informed teachers’ understandings of the terms that they later used to assess student work. Instead of working from existing standards, or descriptions of achievement, the teachers in this study used exemplars of student work to develop sets of bipolar rating scales. These sets represented the constructs that the teachers had identified for visual arts, music, and writing. In writing, for example, the bipolar rating scale “Unevocative (unemotional, distanced)/evocative (of mood or emotion)” (p. 207) was developed to represent the poles of one of the identified constructs. Each of these scales was then converted into a 7-point scale, with a score of 1 representing the beginning of the continuum and a score of 7 representing its end point. The teachers then used these 7-point rating scales to make judgements of student work. Hargreaves et al. (1996) conducted detailed statistical analyses of the teachers’ ratings and found that there was a very high level of agreement between the participants’ judgements of student work. On this basis they concluded that involvement in the development of descriptions of achievement had helped the teachers to clarify their understandings of the terms that they were using to assess student work. This study did not, however, provide any information about the negotiations that the teachers engaged in as they developed the bipolar rating scales that they later used to assess student work.

To understand how taking part in the co-construction of assessment criteria shapes the understandings of achievement at which teachers arrive, attention needs to be paid to the conversations that teachers engage in as they participate in these processes.
Likewise, more research is needed to investigate how using centrally developed standards to inform the development of local criteria shapes teachers’ understandings of standards and criteria. This should provide new insights into whether – and if so, how – participating in criteria-development processes enables teachers to form common understandings of both centrally developed standards and school-developed descriptions of achievement.

Within the National Standards environment, a number of studies (Hipkins & Robertson, 2012; Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015; Wylie & Hodgen, 2010) have provided details about the ways in which schools group teachers for moderation purposes. Given that involvement in moderation is thought to support the development of shared understandings of stated achievement standards (Klenowski & Wyatt-Smith, 2010a, 2014; Maxwell, 2010), inferences can be drawn from these details about the shared understandings that these groupings might have facilitated. In this regard, longitudinal research examining the implementation of the National Standards (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015, 2016) found that schools most often either placed teachers in small groups or had them work as a full staff for moderation purposes. They identified that school size often shaped the way in which schools organised teachers for moderation. In this respect, Ward and Thomas (2012, 2013, 2015, 2016) found that teachers at schools with more than 150 students were most often organised into groups for moderation discussions, while teachers at schools with fewer than 150 students were more likely to work as a full staff. These researchers (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015, 2016) also noted that teachers at some schools were given the opportunity to work in both these configurations.

Hipkins and Roberston (2012) identified a similar pattern in their case study examining the moderation practices of a single primary school. Specifically, teachers at this school were observed participating in moderation discussions at both the syndicate and full-school level. Given that teachers commonly plan for teaching and assessment within their teaching team or syndicate, it is understood that the smaller groupings that were referred to in the longitudinal research (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015, 2016) would also have involved teachers working as a syndicate for moderation purposes. Similarly, Wylie and Hodgen (2010) found that teachers were more likely to engage in moderation discussions with colleagues who taught students at the same year level. If, as Wylie and Hodgen noted, teachers’ moderation activities involve working exclusively or even predominantly with colleagues who teach students
of a similar age, then teachers’ discussions are likely to relate principally to the achievement standards associated with that learning period. This suggests that teachers who are grouped in this way for moderation purposes may only have the opportunity to develop shared understandings of a subset of standards. Accordingly, those teachers who have opportunities to participate in either full-school or both small-group and full-school moderation activities might be more likely to develop shared understandings of a broader range of stated standards.

In addition, it has been noted that some New Zealand primary school teachers have also participated in inter-school moderation activities (Bonne, 2016; Education Review Office, 2010; Hipkins & Robertson, 2012; Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015; Wylie & Bonne, 2014; Wylie & Hodgen, 2010). Presumably, the opportunity to participate in such activities should enable teachers to broaden and reinforce their shared understandings of stated standards. It is, however, unclear whether these inter-school initiatives involved all the teachers from each participating school or just a smaller group. What is clear is that more research is needed to investigate how the ways in which teachers are grouped for moderation purposes affect their development of shared understandings of stated standards.

Given that participation in moderation is expected to enable teachers to develop common understandings of stated standards (Klenowski & Wyatt-Smith, 2010a, 2014; Maxwell, 2010), it follows that their use of these standards during moderation discussions must inform the understandings at which they arrive. Within the National Standards context, it has been identified that New Zealand primary school teachers utilise a range of resources to inform their moderation discussions (Thomas & Ward, 2011; Ward & Thomas, 2013). For example, for those moderation sessions during which writing was a focus, these researchers noted that the resources utilised included the New Zealand Curriculum Reading and Writing Standards (Ministry of Education, 2009c), the National Standards Writing Illustrations (Ministry of Education, n.d.-i), the New Zealand Curriculum (Ministry of Education, 2007), the Literacy Learning Progressions (Ministry of Education, 2010b), the English Language Learning Progressions (Ministry of Education, 2008), and the e-asTTle writing indicators (Ministry of Education, 2012b). In addition, it was also reported that teachers commonly made use of their professional knowledge as well as both school-developed descriptions of performance and annotated work samples.

Although it has been inferred that teachers’ use of these school-developed descriptions of achievement may be contributing to “inconsistencies in the judgements
being made between schools” (Ward & Thomas, 2015, p. 15), very little is known about how this broader range of resources is being used during moderation sessions. In this regard, some critical insights can be gleaned from research conducted in Queensland, Australia. This research (Adie, 2010; Klenowski & Adie, 2009; Klenowski & Wyatt-Smith, 2010b; Wyatt-Smith et al., 2010) examined the social moderation processes associated with the Queensland Curriculum, Assessment, and Reporting (QCAR) trial. With the exception of one study (Adie, 2010), which investigated teachers’ online moderation practices, this research examined the face-to-face moderation processes that teachers engaged in during the QCAR trial. Within this trial, all teachers worked collaboratively to moderate Queensland Comparable Assessment Tasks (QCATs).

These centrally devised assessment tasks were developed to assess students in years 4, 6, and 9 in English, mathematics, and science. Like New Zealand, Queensland has a long history of externally moderated, school-based assessment in the senior phase of schooling. At the time of the QCAR trial, however, standards-referenced moderation processes were being implemented for the first time in years 1–9 (Klenowski & Adie, 2009; Klenowski & Wyatt-Smith, 2010b; Wyatt-Smith et al., 2010).

The teacher moderators who participated in the QCAR trial worked in groups and were provided with three textual resources to guide their work (Adie, 2010; Klenowski & Adie, 2009; Wyatt-Smith et al., 2010). These were a guide, with stated grade descriptors or standards; annotated examples of student work corresponding to each achievement level; and a set of marked student work samples representing each of the possible grades. Because these teachers were working to moderate students’ responses to centrally devised assessment tasks, their use of the stated standards and accompanying resource materials informed their judgement-making processes. It is understood that the teachers’ use of these resources would have shaped the understandings of standards at which they arrived.

Although most groups of teachers involved in the face-to-face moderation meetings attempted to use the resources that had been provided (Klenowski & Adie, 2009; Wyatt-Smith et al., 2010), four different approaches to resource use were identified. For example, some teachers reached agreement about the quality of student assessment evidence with little or no reference to the stated materials that had been provided. These groups of teachers, who were typically relatively experienced, instead based their decisions on unstated personal standards. Although these standards were not discussed, the teachers assumed that they were held in common (Wyatt-Smith et al., 2010). In general, those teachers who did make reference to the stated resource
materials adopted one of three approaches. Some attended almost exclusively to the stated standards (the guide) and made very limited reference to the annotated exemplars. Others paid minimal attention to the stated standards and instead relied on the exemplar material to inform their judgements. Others still used the resources in combination, as had been intended. Although these researchers (Klenowski & Adie, 2009; Wyatt-Smith et al., 2010) did not comment directly upon the effect that the teachers’ varied approaches to resource use had upon the development of shared understandings of standards, Wyatt-Smith et al. (2010) did emphasise the importance of ensuring conceptual clarity in the presentation of such support materials to teachers.

Further insights into the reasons for such variation in resource use can be gleaned from Adie’s (2010) study examining teachers’ online moderation practice. Adie noted that teachers’ attitudes towards the various stated resource materials appeared to shape the ways in which these resources were used. She concluded that a resource’s perceived legitimacy seemed to influence the way teachers used it. In particular, she cited the example of a teacher refusing to utilise the annotated exemplars because these did not align with her personal standards. Because involvement in moderation is expected to enable teachers to develop shared understandings of stated standards, it can be inferred that their use of these standards within moderation sessions must inform the understandings at which they arrive. Yet within the studies associated with the QCAR trial (Adie, 2010; Klenowski & Adie, 2009; Klenowski & Wyatt-Smith, 2010b; Wyatt-Smith et al., 2010), teachers engaged with the stated standards and the accompanying referents in different ways, and in some cases scarcely at all. It seems likely that the varied ways in which these teachers utilised the resource materials would have limited the development of shared understandings of standards within this context.

In comparison with the teachers who participated in the QCAR trial, New Zealand primary school teachers are known to utilise a far greater range of textual resources during their involvement in moderation. Whereas the participants in the QCAR trial made use of just three text-based resources (Adie, 2010; Klenowski & Adie, 2009; Wyatt-Smith et al., 2010), it has been shown that teachers working within the National Standards context can draw on up to seven textual resources during their involvement in moderation (Ward & Thomas, 2013). Further research is needed to identify whether – and if so, how – teachers are able to use their participation in moderation to arrive at shared understandings of stated standards when such a large number of textual resources might be utilised.
The challenges associated with establishing shared understandings of standards have also been alluded to in a report evaluating the arrangements to assure the consistency of teachers’ assessment in Wales at Key Stages 2 and 3 (Estyn, 2010). Following the Welsh Assembly Government’s discontinuation of statutory end-of-key-stage testing for 11- and 14-year-olds in 2005, Wales adopted a system of “teacher assessment” (p. 1). Welsh teachers are required to take part in school- and cluster-based standardisation and moderation processes to support effective teacher assessments and ensure that pupils are assessed accurately and consistently through shared understandings of national standards. Five years after this system was implemented, however, the 2010 evaluation report concluded that although teachers were becoming more confident about their understanding of the national curriculum (NC) level descriptions, differing interpretations of standards continued to persist (Estyn, 2010). Indeed, some schools raised concerns about the difficulty of ensuring that teachers applied “the cluster’s shared understandings of standards” (p. 4) in their subsequent school-based assessments. Interestingly, this report also revealed that in the few local authorities where advisory staff members attended all standardisation and moderation meetings, their presence helped to ensure that the NC level descriptions were applied consistently both within and across clusters of schools. This indicates that it would be worthwhile to explore if and how the involvement of local advisors or experts shapes the understandings of standards that New Zealand primary school teachers attain through their involvement in moderation.

Other research examining how involvement in moderation affects teachers’ understandings of stated standards has focused on participants’ perceptions. Generally, these studies have reported that teachers perceived that their involvement in moderation had helped them either to clarify and extend their understandings of stated standards (Black et al., 2011; Limbrick & Knight, 2005) or to work towards attaining shared understandings of these (Watty et al., 2014). Further research is needed to explore how teachers’ perceptions about the benefits of taking part in moderation correspond to the actual opportunities that these experiences afforded them for developing shared understandings of standards and criteria.

Although some attention has been paid to how involvement in moderation shapes the development of common understandings of standards and criteria, this review of the literature has shown that a number of unanswered questions remain. Specifically, more needs to be known about the ways in which the grouping of teachers for moderation purposes shapes their ability to arrive at common understandings of
standards and criteria. Likewise, further research is needed to ascertain how the range and combination of resources that teachers develop and use to inform their moderation conversations affect the understandings of standards and criteria at which they arrive. Additionally, attention needs to be paid to whether having access to an external expert or advisor shapes the understandings of standards and criteria that teachers develop during their involvement in moderation. Finally, although teachers tend to report that involvement in moderation has supported the development of common understandings of standards, it is unclear whether this is reflected in the actual understandings that are attained. These questions provide the justification for investigating whether – and if so, how – involvement in moderation can provide teachers within the National Standards context with opportunities to clarify and develop shared understandings of stated standards and criteria.

*The quality of assessment evidence*

Within the context of New Zealand’s National Standards, social moderation processes are also expected to improve the dependability of the assessment evidence on which teachers base their judgements of student achievement (Ministry of Education, 2010f). The implication is that teachers will apply the insights that they have gained, through their involvement in moderation processes, to strengthen the quality of the assessment information that they subsequently gather. There has, however, been limited research examining how involvement in moderation shapes teachers’ selection, development, and collection of student assessment evidence. Since the introduction of the National Standards, interest in the assessment evidence that teachers gather for moderation purposes has largely focused on the rationale teachers have employed for selecting student work for moderation sessions (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015, 2016; Wylie & Hodgen, 2010).

Because these studies have examined moderation sessions in which participants discussed their overall teacher judgements (OTJs), much of the focus has been on investigating how teachers selected student assessment information for these purposes. For example, those principals involved in the longitudinal study examining the implementation of the National Standards (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015, 2016) were asked to indicate the selection methods used for choosing work samples associated with the OTJs that their schools had moderated. Specifically, they were asked to indicate whether they had chosen “OTJs near the boundaries between the levels of the standards; The OTJs with inconsistent assessment evidence; A
random selection of OTJs; All OTJs” or some “other” selection approach (Ward & Thomas, 2012, p. 27). Although these data yield some interesting insights into the various ways that schools interpreted the purpose of moderation discussions, they do not allow conclusions to be drawn about what teachers might have learned about the dependability of this assessment evidence from their involvement in these discussions.

In this regard, data about the number of pieces of assessment evidence that teachers used to moderate their OTJs between 2010 and 2012 are of greater assistance (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013). This research showed that the percentage of teachers basing their OTJs of student writing on just one or two pieces of assessment evidence decreased from 29% in 2010 to just 10% in 2012. During the same time frame, the percentage of teachers using between three and eight pieces of evidence for the same purpose increased from 70% to 90%. This shift in the pattern of assessment evidence use is important because it is generally recognised that reliability is affected by the sufficiency of the assessment information that has been gathered (Crooks et al., 1996; Darr, 2005a; J. Smith, 2003). Specifically, Crooks et al. have explained that using “too few tasks” (p. 275) can limit dependability because the size of the sample of student behaviour may not be sufficient for making generalisations about that domain.

On the basis of the longitudinal data that has been cited (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013), it might be inferred that involvement in moderation broadened teachers’ understandings of the benefits associated with using a greater number of sources of evidence to inform their judgements of student achievement. Yet these researchers did not ask the participating teachers to explain why they were collecting more pieces of assessment evidence to inform their judgements of achievement. Because of this it is unclear whether this shift in their assessment practice was associated with their participation in moderation.

A project conducted in the United Kingdom (Harlen, 2010) alluded to a link between teachers’ involvement in moderation and the collection of more dependable assessment evidence. This research indicated that involvement in moderation meetings resulted in teachers collecting and using broader evidence bases to inform their judgements of student achievement. The final phase of this project, which was designed to strengthen teacher summative assessment at the end of Key Stage 2, required participating year 6 teachers to gather assessment evidence over a period of two terms and attend two moderation meetings. By the time of the second moderation meeting, the teachers, who had benefited from the involvement of local authority advisors, were gathering a wider range of assessment evidence. On this basis, it was inferred that these
moderation sessions had facilitated teacher professional learning. It is unclear, however, whether the local authority advisors had played a role in encouraging teachers to bring a larger selection of assessment evidence to these meetings. This indicates that more needs to be known about the contexts and conditions that enable teachers to use their involvement in moderation to improve this aspect of their assessment practice.

The Australian research that was introduced earlier also provides an indication of how involvement in moderation can enable teachers to talk about issues related to the selection and development of assessment tasks. Specifically, Klenowski and Wyatt-Smith (2010b) explained how some participants in the QCAR trial used a moderation session to critique a centrally devised assessment instrument. In this context of engagement and co-construction, a group of teacher moderators sought to mitigate the effects of construct under-representation by making adjustments to the marking criteria. Yet these researchers concluded that there was a “need for greater clarity and explication of construct definition in the task for teachers” (p. 30). This suggests that the changes that these teachers proposed may not have strengthened the assessment task.

Within the National Standards context, teachers can use both school-developed and centrally devised assessment tasks to assess student learning (Ministry of Education, 2010f). Likewise, they can also utilise their own classroom observations as well as work from students’ books to inform their judgements of student achievement. This examination of the literature has indicated that involvement in moderation might provide teachers with opportunities to critically analyse the quality and suitability of the assessment tools and evidence that they use to inform their judgements of student achievement. It has also revealed that very little is known about whether taking part in moderation strengthens this aspect of teachers’ assessment capability. Further research is needed to determine whether involvement in moderation will enable New Zealand primary school teachers to get better at selecting, designing, and administering student assessment tasks.

The quality of teachers’ judgements

Social moderation is presented as a mechanism that can strengthen and align teachers’ judgement making (Harlen, 1994; Klenowski & Wyatt-Smith, 2010a, 2014; Linn, 1993; Maxwell, 2010). This is certainly the expectation within the National Standards context, where moderation is intended to improve the dependability of teachers’ judgements of student achievement (Ministry of Education, 2010e, 2010f). These expectations imply that involvement in moderation will provide teachers with opportunities to learn or
develop the skills required to strengthen their judgement-making processes. Since the introduction of the National Standards, there has been limited research examining how involvement in these processes has affected teachers’ judgement making. For example, although Ward and Thomas (2013, 2015, 2016) raised concerns about the dependability of teachers’ overall judgements of student achievement, their research did not specifically examine how involvement in moderation practice shaped the quality of these judgements. Likewise, although these researchers (Ward & Thomas, 2015, 2016) suggested that teachers’ use of school-developed resources for moderation purposes might have contributed to a lack of dependability, their research did not actually explore this either.

Within the National Standards context, one other study (Wylie & Hodgen, 2010) alluded to a possible link between involvement in moderation and the quality of the judgements at which teachers arrive. Wylie and Hodgen found that teachers who perceived that there was consistency in their school’s OTJs were more likely to work at schools in which OTJs were moderated. That study also identified some variation in the moderation practices of teachers of students in years 1–6 as compared with their counterparts working with students in years 7–8. Specifically, Wylie and Hodgen noted that teachers of students in years 7–8 were less likely to have moderated their OTJs than teachers of students in years 1–6. Within that study, which reports on a national survey of New Zealand primary and intermediate schools, it is understood that teachers of year 7–8 students could have worked at either a full primary (catering for students in years 1–8) or an intermediate school (catering for students in years 7–8). This finding raises the question of whether the type of school (for example, contributing primary, full primary, or intermediate) at which a teacher works could affect how and what he or she learns about assessment through taking part in moderation.

The reporting of teacher perceptions has played an important role in much of the international research examining the relationship between moderation and the dependability of teacher judgements. Numerous studies have noted that involvement in face-to-face moderation sessions increased teachers’ confidence in their ability to arrive at reliable and valid judgements of student work (Black et al., 2011; Kuzich, Groves, O’Hare, & Pelliccione, 2010; Nixon & McClay, 2007; Reid, 2007; Watty et al., 2014). Similarly, Adie (2010) found that having the opportunity to take part in a one-off online moderation meeting had the effect of reassuring a number of participants that their judgements were in line with those of other teachers.
Of these studies, a number also provide some insights into how and why teachers perceived that involvement in moderation had bolstered the confidence they had in their judgement making. For example, one of the teachers involved in the Kuzich et al. (2010) study noted that it was having the opportunity to talk through judgements during moderation sessions that had given him increased confidence that his judgements were “valid and fair” (p. 8). Similarly, another participant in that study emphasised that it was having the opportunity to justify her judgements to others that had been of greatest benefit. Likewise, a teacher involved in the Nixon and McClay (2007) study also commented on the value of having a forum for explicating her judgement-making processes. It is not clear, though, whether taking part in moderation sessions actually enabled the teachers in these studies (Kuzich et al., 2010; Nixon & McClay, 2007) to strengthen their judgement-making processes. Given this, it is possible that their experiences of moderation simply reassured them that their existing approaches to judgement making were sufficient.

In this regard, Adie (2010) found that participating in an online moderation meeting provided a group of teachers with the opportunity to inquire into and strengthen their judgement-making processes. Specifically, these teachers noted how an initial judgement shifted once closer attention was paid to links between evidence in the student’s response and the stated grade descriptors. This experience made a lasting impression and a number of the teachers involved subsequently reported that they had described it to colleagues in their local context. Research like this, which explores how a specific experience of moderation shaped teachers’ understandings of judgement making, is rare. It is valuable because it provides practical insights into how teachers might use their involvement in moderation to strengthen their judgement-making processes. Additional research is needed to examine other ways in which teachers’ experiences of moderation affect their judgement-making processes.

Adie (2010) also noted that being provided with the opportunity to question and examine each other’s judgement-making processes during moderation discussions does not necessarily result in a satisfactory outcome. She cited the example of a group of teacher moderators who, after successfully identifying that they were employing differing rationales for arriving at overall judgements of student work, could not reach agreement about an aggregation approach. Although engaging in this moderation discussion enabled these teachers to detect a factor that certainly would have had an effect on the dependability of their judgement making, an appropriate resolution could not be reached. Further research is needed to identify the circumstances that enable
teachers to use their involvement in moderation to successfully resolve such judgement-making challenges.

Within the National Standards context, only one study has examined teachers’ moderation conversations. That research (Hipkins & Robertson, 2012) concluded that these discussions provided teachers with opportunities to debate a number of issues related to judgement making. One of the issues discussed was the effect that knowledge of the student can have on these processes. This is noteworthy because teacher biases can affect judgement making when assessors draw on knowledge of student abilities or characteristics other than those that are evident in the assessment information (Crooks et al., 1996; Harlen, 2005). Indeed, the moderation of student assessments is expected to limit the effects of teacher biases (Harlen, 2005), and there is an expectation within the National Standards context that involvement in moderation processes will assist with this (Ministry of Education, 2010c). Although participation in moderation enabled the teachers in the Hipkins and Robertson (2012) study to discuss the challenges associated with preventing knowledge of the student from influencing their judgements, it is unclear whether these conversations strengthened their ability to limit the occurrence of this.

Hipkins and Robertson (2012) also reported that involvement in moderation provided teachers with the opportunity to discuss how to score different aspects of students’ writing. Specifically, teachers were observed negotiating how much attention they should pay, in relative terms, to the deep and surface features of student writing. It is understood that these deep and surface features were detailed in the criteria that the teachers used to inform their moderation discussions. These researchers noted that concerns arose about how to weight these attributes when arriving at a judgement. Although few details were provided about how the teachers resolved this issue, involvement in moderation may have provided them with the opportunity to think about how factors associated with the scoring of student performances can affect dependability. In particular, it might be inferred that this experience provided the teachers concerned with the opportunity to consider what Crooks et al. (1996) described as the effect of placing “undue emphasis on some criteria” (p. 272). Crooks et al. explained that taking such an approach could threaten the validity of assessment decisions. Future research needs to attend more closely to whether teachers are able to successfully resolve the judgement-making issues that are identified during moderation sessions. This will enrich current understandings about what and how teachers learn about assessment through taking part in moderation processes.
This section of the literature review has shown that teachers can use their involvement in moderation to identify issues that affect the dependability of judgements of student achievement. The issues acknowledged here included the possible intrusion of teacher biases (Hipkins & Robertson, 2012), inconsistent approaches to the aggregation of assessment information (Adie, 2010), and the assignment of undue weight to certain assessment criteria (Hipkins & Robertson, 2012). What remains unclear, however, is whether teachers are able to use their involvement in moderation to find solutions to the judgement-making issues that they identify. Further research is needed to investigate whether – and if so, how – teachers might also use their involvement in moderation to strengthen their ability to negotiate and resolve such issues. Additionally, although a number of studies have reported that teachers associate involvement in moderation with having increased confidence in their judgement-making ability (Adie, 2010; Black et al., 2011; Connolly et al., 2012; Kuzich et al., 2010; Limbrick & Knight, 2005; Nixon & McClay, 2007; Reid, 2007), only a few have provided evidence of the events or experiences that might actually have enabled them to strengthen this aspect of their assessment capability (Adie, 2010; Hipkins & Robertson, 2012). There is, however, little overlap in the findings from these two areas. Because of this, additional research is needed to explore whether a link exists between the actual opportunities that moderation affords teachers for strengthening their judgement-making ability and the effect that teachers perceive involvement in moderation has on this aspect of their assessment competence.

Assessment for learning

The New Zealand Ministry of Education (2010e) has stated that “moderation supports assessment for learning” (p. 2). Yet only a few studies (Hipkins & Robertson, 2012; Reid, 2007) have investigated the ways in which involvement in moderation enables teachers to utilise or develop their assessment for learning practices. This section of the literature review examines this research and identifies questions that still need to be answered about how taking part in moderation shapes this aspect of teachers’ assessment capability.

One of the tenets of assessment for learning is that a student can only attain a learning goal if she or he understands that goal and can identify what is required in order to achieve it (Broadfoot et al., 2002; Sadler, 1989). In New Zealand, the Ministry of Education has, for a number of years, promoted the practice of teachers developing and sharing learning goals, such as success criteria and learning intentions, with their
students (Ministry of Education, 2009a, 2011a; Poskitt & Taylor, 2008). Yet for these processes to work effectively, teachers require the knowledge and skills to construct well-crafted descriptions of achievement. Research conducted since the introduction of the National Standards has indicated that some teachers are using their involvement in moderation to develop the assessment criteria that they then use to arrive at judgements of student learning (Thomas & Ward, 2011; Wylie & Hodgen, 2010). Although Thomas and Ward (2011) have provided some information about the resources that teachers used to create these criteria, it is unclear whether – and if so, how – taking part in these activities is affecting teachers’ criteria-development skills.

As previously explained, only one study of moderation has examined teachers’ involvement in criteria development. This study (Reid, 2007) was one of several action research projects that were associated with the Scottish Executive Education Department’s (SEED) Assessment is for Learning programme. Although the teachers who participated in this study used a Scottish National Assessment Bank rubric to create a more detailed set of assessment criteria, no conclusions were drawn about whether this experience strengthened their ability to develop assessment criteria. There are numerous sources of advice available to teachers about how to develop rubrics, success criteria, and learning intentions (Brookhart, 2013; Clarke, 2003; Davies & Hill, 2009; Stiggins, 2008). Very little is known, however, about how the experience of working with their colleagues to co-construct such descriptions of achievement shapes teachers’ ability to undertake this work.

In this regard, some attention has been paid to how participating in the co-construction of assessment criteria affects young children’s capability to develop and select these criteria (Higgins, Harris, & Kuehn, 1994). In their yearlong study, Higgins et al. found that providing elementary school students with opportunities to co-construct and select assessment criteria strengthened the students’ ability to engage in these activities. Specifically, they noted that most of the criteria that the students developed at the beginning of the school year were process oriented. By the end of the year, during the course of which the students had ongoing opportunities to work alongside the teacher-researchers discussing and developing assessment criteria, the students tended to identify a balanced mix of process- and product-oriented criteria. Given that there are indications that New Zealand primary school teachers are using their involvement in moderation to co-construct assessment criteria (Thomas & Ward, 2011; Wylie & Hodgen, 2010), it would be useful to examine whether – and if so, how – taking part in these activities is affecting their criteria-development skills.
It is generally recognised that teachers should equip their students with the skills that they need to inquire into and evaluate their own learning (Absolum et al., 2009; Booth, Hill, & Dixon, 2014; Broadfoot et al., 2002; Sadler, 1989). This is because developing the ability to self-assess is thought to enable students to become self-regulating, independent learners. The New Zealand Ministry of Education has suggested that involving students in the moderation process should strengthen their assessment capability (Ministry of Education, 2010e). It has advised teachers that this might be achieved by asking students to select samples of their work that they think exemplify the relevant success criteria. It is not known, however, whether New Zealand primary school teachers are utilising this approach. Likewise, it is unclear at this point whether teachers are using their participation in moderation to identify other ways to involve students in the assessment process.

Only one study (Reid, 2007) has shown that teachers were able to use their involvement in moderation to develop a strategy for strengthening students’ assessment capability. Reid noted that participating in moderation prompted the teachers in her study to work with their students to negotiate assessment rubrics. Because Reid’s study was part of the SEED Assessment is for Learning programme, the participating teachers were provided with the opportunity to familiarise themselves with the literature on formative assessment. Reid argued that their exposure to this knowledge base, coupled with the confidence that the teachers had gained from their earlier involvement in criteria development, prompted them first to share these criteria with their students and then to involve them in the co-construction of assessment rubrics. Further research is needed to identify whether taking part in moderation processes might enable teachers to develop other strategies for strengthening students’ assessment capability.

The New Zealand Ministry of Education (2010e) has suggested that the conversations that teachers participate in during moderation sessions should enable them to adjust their teaching and learning programmes “to meet student learning needs” (p. 2). This is one of the ways in which the Ministry of Education has identified that involvement in moderation should support assessment for learning. One New Zealand study (Hipkins & Robertson, 2012) noted that teachers used their post-moderation discussions to identify next learning steps. These researchers explained that it was during this phase “that the teachers became directly accountable to their students” (p. 48). Although this research linked teachers’ involvement in moderation with their subsequent identification of next learning steps, it did not examine whether taking part in these moderation conversations shaped the teachers’ ability to inquire into and adjust
their teaching and learning programmes. One other study (Reid, 2007) reported that teachers felt that involvement in moderation had improved their ability “to determine teaching foci for writing lessons” (p. 141). Yet this seemed to be linked with the teachers’ participation in the co-construction of assessment criteria rather than their involvement in conversations about student work. Further research is needed to investigate whether having the opportunity to discuss judgements of student work with their colleagues shapes teachers’ ability to inquire into and adjust their teaching and learning programmes.

Very little attention has been paid to the ways in which involvement in moderation shapes teachers’ assessment for learning capability. Although a few studies (e.g., Hipkins & Robertson, 2012; Reid, 2007) have identified ways in which teachers integrated assessment for learning principles and practices into their moderation processes, additional research is needed to identify whether teachers might use their participation in moderation to strengthen this facet of their assessment capability.

Social moderation and teacher learning

A number of studies (Adie, 2010; Hipkins & Robertson, 2012; Klenowski & Adie, 2009; Reid, 2007; Wyatt-Smith et al., 2010) have used Wenger’s (1998) social theory of learning to investigate how involvement in social moderation informs teacher learning. This is fitting, because social moderation is an inherently participative, consultative process (Maxwell, 2010). As its name suggests, this form of moderation is reliant upon social participation. Likewise, Wenger (1998) has conceptualised learning as a socially participative process. He has identified four components that characterise learning as social participation: practice (learning as doing), meaning (learning as experience), identity (learning as becoming), and community (learning as belonging). As Wenger has conceived it, social participation is an encompassing process that involves actively taking part in the practices of a social community, or what he has termed a “community of practice” (p. 5). For Wenger, a community gains coherence as its participants, or members, make use of a shared repertoire to mutually engage in a joint enterprise. Wenger has posited that participation in a community of practice shapes not only what members do, but also who they are and how they interpret what they do.

Given this, Wenger’s (1998) conceptual framework offers a mechanism for investigating how teachers’ mutual engagement in the joint enterprise of moderation practice informed the understandings about assessment at which they arrived. Because the term social in social constructionism relates to the mode of meaning making (Crotty,
Wenger’s (1998) social theory of learning is compatible with the social constructionist perspective that underpins this thesis. For these reasons, Wenger’s conceptual framework is employed here to examine how and what teachers learned about assessment through their involvement in social moderation. More specifically, this thesis uses Wenger’s conception of practice to investigate how participation in moderation informed teachers’ understandings of assessment for learning principles and practices. Likewise, his characterisation of meaning is drawn on to examine how involvement in moderation shaped teachers’ understandings of those factors that affect the dependability of assessment information. Additionally, because Wenger has linked his conception of identity with perceptions of competence, this notion is used to examine how involvement in moderation processes informed teachers’ perceptions of their assessment capability. This approach should provide insights into the ways in which involvement in moderation contributed to each of these facets of teachers’ assessment capability. Likewise, it should also enable conclusions to be drawn about whether taking part in moderation could facilitate the formation of assessment-focused communities of practice.

Wenger (1998) has, however, sounded a note of caution against idealising the role that communities of practice can play in relation to learning. In this regard he has explained that “the indigenous production of practice makes communities of practice the locus of creative achievements and the locus of inbred failures” (p. 85). Interestingly, Timperley et al. (2007) arrived at a similar conclusion. These researchers systematically reviewed the evidence-based research on teacher professional learning and development and identified a range of characteristics associated with the promotion of effective professional learning opportunities. In particular, they found that involvement in what they referred to as “a professional community of practice” (p. xxx) did not on its own necessarily provide teachers with effective professional learning opportunities. Indeed, although they noted that involvement in such a community appeared to be a necessary condition for effective professional learning, they also found that these communities could reinforce ineffective practices.

Moreover, Timperley et al. (2007) concluded that a range of other contexts also shaped the efficacy of teacher professional learning opportunities. These included, but were not limited to, the provision of extended time for opportunities to learn, the involvement of an external expert, the existence of active school leadership, the provision of a rationale for engagement in learning, and the adoption of approaches that were consistent with wider trends in policy and research. In fact, in earlier research I
pointed to the existence of a relationship between the access that teacher moderators had to a number of these learning contexts and the learning opportunities that involvement in moderation afforded them (Smaill, 2013). Given that these contexts appear to offer a complementary framework within which to examine the professional learning opportunities that involvement in moderation provides, the role that they play in the moderation processes associated with New Zealand’s National Standards assessment system will also be explored in this thesis.

**Research questions**

To focus this inquiry, the preceding literature review has informed the development of a series of subquestions. These questions draw upon Wenger’s (1998) conception of learning as a socially participative process that involves doing (practice), experiencing (meaning), becoming (identity), and belonging (community). Specifically, these questions seek to explore how involvement in the social moderation practices associated with the National Standards assessment system shape the meanings or understanding about assessment that teachers generate and the assessment identities that they assume. This should enable conclusions to be drawn about whether taking part in these moderation activities could enable New Zealand primary school teachers to participate in assessment-focused communities of practice. Therefore, the subquestions for this thesis are:

1. How and in what ways does involvement in social moderation shape teachers’ assessment for learning capability?
2. How and in what ways does involvement in social moderation shape teachers’ understandings of factors that can affect the dependability of assessment information?
3. How and in what ways does involvement in social moderation shape teachers’ perceptions of their assessment capability?
4. Can involvement in social moderation be used to support the formation of assessment-focused communities of practice, and if so, how?
Chapter 3. Methodology

This chapter describes the ways in which social constructionism (Crotty, 1998) was used in conjunction with Wenger’s (1998) social theory of learning to inform the methods that were employed within this study. It explains how each of these understandings shaped the ways in which the data were both produced and analysed. First, it explains why a case-based approach (Stake, 2006; Yin, 2014) was used for producing information about the learning experiences that involvement in moderation afforded the teachers at three schools. Next, a description of the processes involved in selecting and approaching these schools is presented and a brief introduction to the key participants at each school is offered. Following this, the methods that were used to produce data are explained and a rationale for their inclusion is outlined. A description of the analytic approach, which is iterative and integrates key elements of grounded theory (Charmaz, 2006), is then presented. After outlining the ethical considerations associated with this study, the chapter concludes by reflecting upon how this inquiry might have shaped the social moderation activities that it set out to examine.

The use of case

Wenger’s (1998) social theory of learning is premised on the understanding that learning is a product of active participation in the practices of a social community. Drawing on the work of both Yin (2014) and Stake (2006), this study utilises a multi-case approach to investigate how and what the teachers at three schools learned about assessment through their active participation in social moderation activities. Case study research is ideally suited to inquiries that investigate a contemporary phenomenon within its real-world context (Yin, 2014). Here, social moderation practice is the contemporary phenomenon of interest or the “quaintain” (Stake, 2006, p. 6), with the New Zealand year 1–8 primary school setting providing the context. To explore how social moderation practices are enacted in a variety of real-world contexts and to examine the learning opportunities associated with active participation across these different contexts, three cases were chosen. Each case, which takes the form of a school, functions as a “specific, unique, bounded system” (Stake, 1994, p. 237). Within this study, the potential similarities and differences between cases were important for what

2 The term data production (as opposed to data collection) is employed within this thesis to reflect the social constructionist perspective (Crotty, 1998) that underpins this research.
they could reveal about the phenomenon. Therefore, my interest in each case was primarily “instrumental” (Stake, 2006, p. 8). Given this, and to protect the anonymity of participants and schools, no individual case reports are presented. Instead, using an approach that Yin (2014) has described, each of the findings chapters explores a specific cross-case issue. Within each of these chapters, information from the individual cases is provided to illuminate and substantiate the findings.

In keeping with the social constructionist perspective (Crotty, 1998) that underpins this study, using a case-based approach allowed me to produce or construct meaning about social moderation through my interaction with the research participants (Hamilton & Corbett-Whittier, 2012). This study’s findings are, therefore, based upon my interpretations of the participants’ lived experiences of this phenomenon. The upcoming chapter details the steps that were taken to ensure the trustworthiness of these findings. Throughout this chapter and the remainder of this thesis, pseudonyms are used in the discussion of schools and teachers to protect the identity of the participants.

Selecting and gaining access to schools

Yin (2014) explained that the case study method is particularly well suited to those inquiries in which the “boundaries between phenomenon and context may not be clearly evident” (p. 16). To examine if and how a series of contextual factors shaped the phenomenon of interest, the sampling process for schools was purposive. Through my review of the literature on the early implementation of the National Standards, I identified three contextual factors that I thought might have an effect on social moderation. These were school size, school type, and assessment preparedness. Specifically, Ward and Thomas (2012, 2013, 2015, 2016) found that school size often dictated the way in which schools organised teachers for moderation purposes. They identified that schools with more than 150 students typically grouped teachers for moderation discussions, while schools with fewer than 150 students generally worked as a full staff. Because I was curious about whether the grouping of teachers for moderation purposes would affect learning opportunities, I sought to work with schools with both fewer and more than 150 pupils. Additionally, because Wylie and Hodgen (2010) identified some variation in the moderation and assessment practices of teachers of students in years 1–6 as compared with their counterparts working with students in years 7–8, school type emerged as a factor of interest. Finally, given that an Education Review Office report (2010) hinted at a possible relationship between a school’s assessment preparedness and their capacity to develop and engage in social moderation
activities, I was also interested in how assessment preparedness might affect the learning opportunities that involvement in social moderation would afford participants.

To identify schools that would enable me to examine social moderation in relation to each of these contexts, I drew on insider knowledge that I had attained through working with local schools. In the years before commencing this study I had worked with schools in my region, first as a primary school teacher and later as a researcher on a university-based longitudinal study of student achievement. These experiences meant that I had a broad knowledge of the schools within my local area that catered for students in years 1–8. To deepen this understanding, I also turned to the Education Review Office’s (ERO) website (New Zealand Government, n.d.), where I read current reports of schools’ reviews. These reports contain basic information about the quality of a school’s assessment systems and their assessment preparedness. Finally, to glean up-to-date information about both school size and type, I accessed a database of information about New Zealand schools that is publicly available on a New Zealand government website (Education Counts, n.d.). After reviewing this information, I selected three schools that, as a group, seemed to offer both “potential for learning and representativeness” (Stake, 2006, p. 25). By this I mean that I selected schools that I thought would be participating in moderation activities during the data-collection phase, have an interest in participating in the study, and afford insights into the three identified contexts.

Within my initial correspondence with principals, I briefly outlined the purpose of my study and requested an opportunity to meet with each of them individually. The purpose of the proposed meeting was to obtain feedback about the study’s research design and to discuss whether their school might be interested in participating. This collaborative approach reflected my commitment to the principles of fairness (Guba & Lincoln, 1989) and reciprocity (Lincoln, 2009). I wanted to ensure that potential participants had the opportunity to identify issues that they felt needed investigating in relation to social moderation. I also felt that it was important to involve the principals in decisions concerning the best way to go about exploring these issues. I hoped that this consultative approach would make the study more useful and relevant for prospective participants and therefore increase the likelihood that I gained access to each school.

Researchers sometimes experience problems gaining access to sites. For example, access is at times denied when sites or participants have concerns about exposing “quasi-private worlds to public scrutiny” (Heath, Charles, Crow, & Wiles, 2007, p. 410). Involvement in moderation requires teachers to de-privatise their
judgement-making processes (Klenowski & Adie, 2009; Wyatt-Smith & Bridges, 2006). Given that the introduction of the National Standards (Ministry of Education, 2009c) had required New Zealand primary schools to develop new moderation processes (Ministry of Education, 2010e), I was concerned that principals and teachers might feel particularly uneasy about allowing a researcher to observe these fledgling activities. I was also aware that researchers are sometimes refused access to a site if participants are already busy or have existing demands on their time (Basit, 2010; Heath et al., 2007). I felt that this was of particular concern because I knew firsthand how time-poor teachers often are. Moreover, even without the burden that can be associated with participating in a study, moderation is generally perceived as a time-consuming activity (Gipps, 1994; Klenowski & Adie, 2009; McNaughton, 2009).

As had been intended, these meetings provided principals, and in one case also a deputy principal, with the opportunity to share their views about the study’s design. Their feedback led me to reduce the number of teacher interviews and to audio-record rather than video-record the moderation meetings. At that point, all three principals provided consent for their schools to participate in the yearlong study. The option of adding a fourth school was considered. This, however, was discounted when it became clear that the likely scheduling clashes among after-school meeting sessions at the various schools would outweigh the benefits associated with involving another school. Specifically, because schools typically assign Monday, Tuesday, or Wednesday afternoons for full-staff and syndicate meetings, I anticipated that adding a fourth school would increase the likelihood that two of the participating schools would schedule meetings at the same time. It was also clear that the three schools that I received permission to work with fulfilled Stake’s (2006) requirement of affording sufficient opportunity to examine how the phenomenon of interest performed in different environments. These differences are presented in Table 1 in a manner that protects the anonymity of each school.
Table 1

*Contextual characteristics of cases in the study*

<table>
<thead>
<tr>
<th>Context</th>
<th>Description</th>
<th>Number of schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>School size</td>
<td>Fewer than 150 students</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>More than 150 students</td>
<td>2</td>
</tr>
<tr>
<td>School type</td>
<td>Contributing primary school: Caters for students in years 1–6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Full primary school: Caters for students in years 1–8</td>
<td>1</td>
</tr>
<tr>
<td>Assessment systems</td>
<td>Assessment systems identified as an area for improvement in recent ERO report</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Assessment systems identified as an area of strength in recent ERO report</td>
<td>2</td>
</tr>
</tbody>
</table>

**Participants**

At each school, all classroom teachers had some involvement in the study and were invited to assume one of three participant roles. These roles were lead teacher, focus-syndicate teacher, and focus-school teacher. The method for selecting the lead teacher at each school was purposive. Other participant roles were then assigned depending on the working relationship that existed with the lead teacher. To select the lead teacher for each school, principals were asked to identify the teacher who had assumed responsibility (independent of this study) for organising and leading moderation sessions. This role was created for two reasons. First it was intended to provide a liaison for me regarding organisational matters relating to moderation. Second, it was hoped that developing a relationship with each school’s lead teacher would afford insights into the decisions that they had made regarding the organisation of their school’s moderation activities.

Based on my preliminary conversations with each principal, I knew that most moderation activities would occur during syndicate and full-staff meeting sessions. Therefore, to develop understandings about how and what teachers learn about assessment through participating in social moderation activities, I needed to attend as many of these meetings as possible. In New Zealand, both contributing and full primary schools (henceforth, primary schools) are typically organised in syndicates or teaching...
teams. Depending on the size of a school there are either two or three syndicates, each of which has an assigned leader. Because each school in my sample scheduled a common time slot for their syndicate meetings, it was not possible for me to attend all syndicate meetings at a given school. Instead, I chose to form a close relationship with one syndicate at each school. For practical reasons, I selected the syndicate within which each lead teacher taught. At each school, I called this the focus syndicate and described the other members of this group as focus-syndicate teachers.

As well as consenting for me to observe and audio-record all the moderation meetings in which they were involved, focus-syndicate teachers also agreed to take part in either an individual or a group interview and to complete a questionnaire. To enable me to observe full-staff moderation sessions, I asked the principal and the remaining teachers at each school for permission to observe and audio-record these meetings. The teachers in this group (whom I referred to as focus-school teachers) also agreed to complete the questionnaire. Similarly, the principals consented to take part in an interview. Samples of all the information sheets and consent forms that were associated with this study are included in Appendix A. Because the members of each school’s focus syndicate played a particularly important role in this study, these participants are now introduced. To preface this, Table 2 provides additional details about each focus syndicate.

Table 2

<table>
<thead>
<tr>
<th>School</th>
<th>Syndicate type</th>
<th>Year levels of students taught</th>
<th>Participating teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Heights</td>
<td>Senior</td>
<td>3–6</td>
<td>1 lead teacher</td>
</tr>
<tr>
<td>Greenville School</td>
<td>Middle</td>
<td>2–4</td>
<td>1 lead teacher</td>
</tr>
<tr>
<td>Riverside School</td>
<td>Senior</td>
<td>5–8</td>
<td>1 lead teacher</td>
</tr>
</tbody>
</table>
Central Heights School

Kerry led the senior syndicate and held the position of deputy principal at Central Heights School. Kerry was lead teacher for this study because she had existing responsibility for organising her school’s moderation activities. The data-collection phase coincided with Kerry’s 30th year of teaching and her 15th year of working at Central Heights School. During the data-collection period she taught a class of year 5–6 students. She had extensive experience teaching at this level.

Margaret was in her 25th year of teaching and it was the 14th year she had taught at Central Heights School. As the teacher of a year 3–4 class, Margaret taught the youngest students in the focus syndicate. This was the 10th year that she had worked with students in these year levels.

William, who taught a year 4–5 class, had a close working relationship with Margaret. At the time of data collection, William was in his 7th year at Central Heights School. This was his 14th year working as a teacher and his 7th year teaching students in years 4–5.

Chris had first taught at Central Heights School as a trainee teacher. During this time he had worked in Kerry’s classroom. The data-collection phase coincided with his 10th year of teaching and his 5th year working at Central Heights School. Because he was teaching a year 5 class at this time, Chris often worked closely with Kerry.

Greenville School

Susan was responsible for leading the middle syndicate at Greenville School, where she taught a year 3–4 class. Although Susan had over 14 years of teaching experience, it was only the 2nd year that she had worked with students at these year levels. Susan was the lead teacher for this study. The principal explained that he had suggested Susan for this role because she had independently approached him to discuss updating Greenville School’s moderation processes. This was her 6th year at Greenville School.

Sophia had started teaching at Greenville School 5 years earlier as a beginning teacher. During this time she had worked mostly with students in years 2 and 3. These were the year levels that she was teaching during the data-collection period.

Phillip had taught for over 5 years in the United Kingdom before moving to New Zealand. At the time of data collection he was in his 3rd year at Greenville School,
where he was teaching a class of year 4 students. It was the 4th year he had taught
students at this year level.

**Riverside School**

Stephen led the senior syndicate at Riverside School, where he also held the position of
deputy principal. At the time the study was conducted, Stephen had been teaching for
almost 30 years and it was his 5th year at Riverside School. He taught a year 6–7 class.
It was the 15th year he had taught students in these year levels. Although the principal
of Riverside School stated that she would take primary responsibility for organising this
school’s moderation activities, she asked Stephen to become the lead teacher for this
study because he had independently expressed an interest in moderation.

Kim was in her 11th year of teaching, having spent all but her 1st year teaching at
Riverside School. Kim taught a class of year 7 and 8 students. It was the 3rd year she
had worked with students at these year levels.

Abby taught a class of year 5 and 6 students. She explained that it was the 10th year she
had worked with students in these year levels. She was in her 15th year of teaching and
it was her 12th year at Riverside School.

**Data production**

Within this study, I sought to produce data that would provide insights into how and
what teachers had learned about assessment through their involvement in social
moderation activities. According to Wenger (1998), learning occurs in such contexts
through doing, experiencing, becoming, and belonging. To understand how teachers
learned through engaging in these processes, the research involved the production of
data generated through observations and recordings of moderation meetings, semi-
structured interviews, and a teacher questionnaire. To ensure that multiple perspectives
were obtained (Hamilton & Corbett-Whittier, 2012; Yin, 2014), interviews were
conducted with both teachers and principals. Observation and interview sessions also
provided an opportunity to produce field notes and gather artefacts. A series of Likert-
type questions in the teacher questionnaire produced a small amount of quantitative data.
These quantitative data were used to complement the interview and observation data.
The flexibility that case study research affords for collecting data from multiple sources
is one of its strengths (Hamilton & Corbett-Whittier, 2012; Yin, 2014). Using a variety
of methods for producing data means that it is possible to build on the strengths of each
approach by ensuring a complementary method addresses, and thereby mitigates, weaknesses (Patton, 2002). Moreover, the use of a variety of data sources also allows findings to be triangulated (Denzin & Lincoln, 2000). This lends depth and rigour to the processes of data production and analysis, and reinforces the legitimacy of the conclusions that can be drawn.

A summary of the observation, interview, and questionnaire data that were produced across the three schools is presented in Table 3. As explained earlier, it was only the teachers within each school’s focus syndicate who participated in an interview. Because some teachers preferred to take part in a paired or group interview, this table shows the number of teacher interviews that were conducted as well as the total number of teachers who participated in interviews.

Table 3

<table>
<thead>
<tr>
<th>Type and amount of data collected for each school (excluding artefacts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data type</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Moderation meetings attended</td>
</tr>
<tr>
<td>Teacher interviews</td>
</tr>
<tr>
<td>Teachers interviewed</td>
</tr>
<tr>
<td>Principal interviews</td>
</tr>
<tr>
<td>Teacher questionnaires</td>
</tr>
</tbody>
</table>

Researcher stance and identity

The observation of moderation meetings played a very important role in this study. To produce these data, I adopted an observer-as-participant stance. In this regard, my interpretation of this role exceeded Gold’s (1958) classic definition of the observer as participant. According to Gold, this role is used in studies that involve “one-visit interviews” (p. 221) and therefore entail very little observation. Instead, I assumed a stance consistent with Merriam’s (2009) description of the observer as participant. She defined this stance as one in which the researcher makes her or his activities known to the group, while giving greater priority to information gathering than group
participation. In my case, although I elected to participate a little in each school’s moderation sessions, I did not take part in those tasks that comprised the group’s core activities (e.g., criteria development, task development, and judgement making). I assumed this stance for two reasons.

First, although I was aware that the experience of being observed would have an effect on the participating teachers (Merriam, 2009), I wanted to minimise this. Because I anticipated that each school’s interpretation and enactment of moderation would shape what teachers learned through their involvement in these activities, I did not want my participation to affect the approach to moderation that was adopted. This influenced my decision not to participate in the core moderation activities at any of the schools. Second, I understood that my participation in moderation meetings would shape the membership role that I negotiated at each school (Angrosino & Rosenberg, 2011). In this regard, I sought to negotiate what Adler and Adler (1987) described as “peripheral membership” (p. 36). Being accorded peripheral membership can provide researchers with an insider status or view, while also allowing them to maintain an outsider perspective. Patton (2002) stated that achieving this dual perspective can enable the researcher not only to understand the setting as an insider but also to describe it to and for outsiders.

As I sought to negotiate a membership role at each of the schools, I drew upon my lived experiences as a teacher and as a researcher in the field of educational assessment. At each school, I introduced myself as a fully registered primary school teacher. In this way, I presented myself as someone with lived experiences of teaching and assessing writing in school contexts that were very similar to their own. Morse (2011) explained that being aware of the “currency” (p. 410) of a site can enable a researcher to build relationships with participants that support the generation of rich data. My teaching background meant that I had firsthand experience of attending lengthy staff meetings after a full day in the classroom. Moreover, I recalled how much I appreciated being provided with a snack on such occasions. In an effort to foster trusting, collegial relationships with teachers, I decided to bring baking to each of the meetings that I attended. The teachers appreciated this and their expressions of gratitude are captured in the transcripts of most observations. In time, and perhaps partly because of my efforts to participate in and contribute to moderation meetings, I felt that I was granted insider status at each of the schools. Because the participating teachers appeared to feel comfortable speaking frankly and openly about a wide range of topics and issues in my presence, I believe that this insider status contributed to the richness of the data that were produced.
Yet the process of negotiating peripheral membership at each of the schools also introduced some challenges. Because the teachers knew that I worked in the field of educational assessment, I was at times asked to provide guidance and answer questions about assessment issues. This occurred at all schools. Typically, the questions that I fielded were of a minor nature. For example, on one occasion a participant who knew that I was familiar with a specific standardised test asked me to clarify its marking procedure. Because it seemed important to respond to such questions, I included a systematic record of all the advice that I provided to schools in my field notes. This enabled me to account for any impact that my responses to these requests might have had on both the data that I obtained and my interpretations of those data (Merriam, 2009). This issue is discussed in further detail later in the chapter.

Observation of moderation meetings

I selected observation as a data-production method for the access that this approach would provide to the phenomenon of interest: social moderation. The opportunity to observe teachers as they engaged in moderation meetings provided me with numerous firsthand encounters with teachers’ moderation activities. As previously stated, Wenger (1998) has used four interrelated components to characterise his social theory of learning. These are practice, meaning, identity, and community. Through the observation of moderation meetings, I hoped to produce data that would provide insights into how and what teachers learned about assessment in relation to each of these components. For example, Wenger defined the concept of practice as “a way of talking about the shared historical and social resources, frameworks, and perspectives that can sustain mutual engagement in action” (p. 5). Observation of moderation meetings potentially offered a way of gathering information about the tools, systems, perceptions, and procedures that the teachers employed to sustain their mutual engagement in this work. I recognised that the observation of moderation meetings might render limited information about the actual meanings and understandings that participants attained during these sessions (Patton, 2002). Yet I hoped that the participative, generative nature of moderation practice might mean that the observations yielded some insights in these regards.

A total of 21 moderation meetings were observed across the three schools. Three types of data were produced during these observations. These were audio recordings, field notes, and artefacts. With the consent of all participants, digital recording devices were used to capture teachers’ discussions during moderation meetings. In most cases,
these recordings later enabled me to transcribe teachers’ conversations in full. There were, however, three exceptions to this. These all occurred when teachers conducted some of their discussions in smaller subgroups. This resulted in a series of simultaneous conversations that the recording equipment rendered unintelligible. After identifying this problem, I employed a number of remedial strategies. During the final meeting in which teachers worked in subgroups, the use of a set of directional microphones later enabled me to fully transcribe the conversation of one subgroup of teachers.

Across the three schools, teachers were observed developing and/or making use of a range of resources and artefacts during their moderation work. These included the centrally developed resources (Ministry of Education, 2003b, 2009c, 2010b, 2012b) that were introduced in the previous chapter, locally developed assessment resources (see Appendices B & C), pedagogical texts (Department for Education and Employment, 2000; Loane & Muir, 2010; A. Smith, 2009; Wing Jan, 2001), and in one case a series of National Education Monitoring Project (NEMP) assessment tasks (Crooks, Flockton, & White, 2007). When teachers created or made use of a locally developed resource, I asked to be provided with a copy of it. I also independently sourced any pedagogical texts that teachers referred to during moderation meetings.

At each school, I also used a journal to record field notes related to each observation session. These notes (see Appendix D) were produced during and immediately following each of the observed meetings. Within these notes, I logged information that an audio recording could not capture. This included details about the meeting setting, information about the seating positions of participants, and instances when participants referred to artefacts to guide their conversations. To assist with transcription, I also attempted to signal the main conversational turns that occurred during each meeting. I marked each of these with a time stamp and either the speaker’s name or (during initial meetings at each school) a description of the speaker. After leaving a meeting, I also noted any questions that the observation had raised and identified issues that I needed to follow up on during subsequent observations and analyses. Finally, when required, these field notes also included my reflections about the possible effect of my presence on a meeting. In this way, these notes became an important log of my research progress and constituted part of my “audit trail” (Merriam, 2009, p. 222). Because of this, the construction of these field notes helped lend trustworthiness to my findings. A complete record of the data that were gathered and produced from the observation of moderation meetings is presented in Table 4.
Table 4

*Summary of data resulting from the observation of moderation meetings*

<table>
<thead>
<tr>
<th>Data type</th>
<th>Central Heights</th>
<th>Greenville</th>
<th>Riverside</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation transcripts</td>
<td>Meetings 1–12</td>
<td>Meetings 1–6</td>
<td>Meetings 1–4</td>
</tr>
<tr>
<td>Observation field notes</td>
<td>Meetings 1–12</td>
<td>Meetings 1–6</td>
<td>Meetings 1–4</td>
</tr>
<tr>
<td>Overview of meetings</td>
<td>See Appendix E</td>
<td>See Appendix F</td>
<td>See Appendix G</td>
</tr>
<tr>
<td>School-developed artefacts</td>
<td>Success Criteria: Personal Experience</td>
<td>Must, Should, Could Chart: Year 2</td>
<td>Must, Should, Could Chart: Year 2</td>
</tr>
<tr>
<td></td>
<td>Success Criteria: Report &amp; Recount</td>
<td>Must, Should, Could Chart: Year 3</td>
<td>Must, Should, Could Chart: Year 3</td>
</tr>
<tr>
<td></td>
<td>Success Criteria: Character</td>
<td>Must, Should, Could Chart: Year 4</td>
<td>Must, Should, Could Chart: Year 4</td>
</tr>
<tr>
<td></td>
<td>Success Criteria: Narrative</td>
<td>Must, Should, Could Chart: Year 5</td>
<td>Must, Should, Could Chart: Year 5</td>
</tr>
<tr>
<td></td>
<td>School moderation plan</td>
<td>Tracking Sheet: Year 3</td>
<td>Tracking Sheet: Year 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WALT &amp; WYLF Chart: Recount</td>
<td>WALT &amp; WYLF Chart: Recount</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WALT &amp; WYLF Chart: Description</td>
<td>WALT &amp; WYLF Chart: Description</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WALT &amp; WYLF Chart: Narrative</td>
<td>WALT &amp; WYLF Chart: Narrative</td>
</tr>
<tr>
<td>Data type</td>
<td>Central Heights</td>
<td>Greenville</td>
<td>Riverside</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Other artefacts</td>
<td>I’ve got something to say: Leading young writers to authorship (Loane &amp; Muir, 2010)</td>
<td>The national literacy strategy: Framework for teaching (Department for Education and Employment, 2000)</td>
<td>NEMP writing assessment task: How to get to Ani’s place (Crooks et al., 2007, p. 37)</td>
</tr>
<tr>
<td></td>
<td>Write ways: Modelling writing forms (Wing Jan, 2001)</td>
<td>PM writing: Exemplars for teaching writing (A. Smith, 2009)</td>
<td>NEMP writing assessment task: Popcorn (Crooks et al., 2007, p. 43)</td>
</tr>
<tr>
<td></td>
<td>PM writing: Exemplars for teaching writing (A. Smith, 2009)</td>
<td></td>
<td>NEMP writing assessment task: Torch (Crooks et al., 2007, p. 38)</td>
</tr>
<tr>
<td></td>
<td>e-asTTle writing seminar slides (Evaluation Associates, 2012)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: WALT = We Are Learning To; WYLF = What You’re Looking For*
Interviewing participants can provide the researcher with the opportunity to produce information about those things that cannot be directly observed (Forsey, 2012; Patton, 2002; Stake, 1994). Within this study, the primary motivation for conducting interviews was to gain insights into those aspects of the participants’ learning that could not be observed during the moderation meetings. In this way, the combination of interviews and observations was complementary. In particular, interviews seemed to offer a means of exploring how involvement in moderation activities had enabled participants to negotiate meanings about assessment and develop their identities as assessors. For these reasons, I understood that conducting interviews would enable me to examine those dimensions of the participants’ learning that Wenger (1998) has referred to as meaning and identity. An additional goal of the interviews was to gather information about historic events or opportunities that may have shaped each school’s moderation practice.

Within this study, a semi-structured or “interview guide approach” (Patton, 2002, p. 349) was employed to produce these data. Such an approach enables the same topics and issues to be canvassed in each interview while allowing the exact sequence and wording of questions to be determined in response to the given situation (Merriam, 2009; Patton, 2002). This approach was chosen because it afforded me the flexibility I needed to respond to the unique ways in which moderation had been interpreted and enacted at each school. Two interview guides were developed to inform this process: one for use with principals (see Appendix H) and the other for use with both focus-syndicate and lead teachers (see Appendix I). To elicit information about how lead teachers had experienced the process of planning, organising, and facilitating these activities, this group of participants was asked questions about some additional topics (see Appendix J). All interviews were conducted late in the data-collection phase and took place in either the third or fourth term of the school year. This provided participants with the opportunity to share their perspectives about the full range of moderation activities in which they had been involved. It also enabled me to use what I had observed during moderation sessions to inform how I presented the questions.

From a social constructionist perspective, the interview is a social setting or a shared space in which both the interviewer and the interviewee(s) co-construct data (Roulston, 2010). From this perspective, an ideal interview is one in which both the interviewer and the interviewee(s) share the interview space equally (Gubrium & Korol-Ljungberg, 2005). This, however, can be hard to achieve. For example, Gubrium and
Koro-Ljungberg noted that the control the researcher has over the topics canvassed during an interview and the interpretations that are ultimately reached during analysis often contributes to a power imbalance. In an effort to reduce this imbalance, I adopted a reflexive stance when planning and conducting participant interviews. This involved sharing interview topics with participants in advance, inviting participants to select the interview time and location, adopting a relatively flexible approach to questioning, providing participants with the opportunity to review and edit interview transcripts, and agreeing to group rather than individual interviews when participants expressed a preference for this. With regard to this final point, Table 5 provides an overview of how teachers and principals participated in these interviews.

Table 5

*Summary of teacher and principal interview participation*

<table>
<thead>
<tr>
<th>Interview configuration</th>
<th>Lead teacher</th>
<th>Focus-syndicate teacher</th>
<th>Principal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Group</td>
<td>1</td>
<td>4</td>
<td>—</td>
</tr>
</tbody>
</table>

Although the decision to include group interviews introduced some challenges, it also yielded a range of benefits. For example, at times, the conversations that occurred during group interviews went in unanticipated directions. This meant that I had to make decisions about how to ensure these conversations flowed naturally without having them stray too far from the parameters of the interview guide. As Currie and Kelly (2012) explained, however, conversations are more likely to go in unanticipated directions within a group interview setting because the group dynamic increases the level of control that the participants have over the direction of the conversation. Given this, I understood that these unanticipated conversational turns provided evidence that the group interviews served to flatten the researcher-participant hierarchy. Another benefit associated with group interviews is that the participant interaction they allow can result in the production of particularly rich data (Currie & Kelly, 2012). An example of this can be seen in the following exchange, during which two teachers responded to a question about the primary purposes of moderation:
Margaret: Yeah, well I think that it’s [the primary purpose is] just the
Margaret and William: consistency
Margaret: really is the big thing isn’t it. And also support.
Because in getting to that point of consistency we actually
do support each other and I think we, you know, you hear
other people saying something and you think, oh, that’s a
good idea, I could do that.
William: We learn from each other.
Margaret: So we do learn from each other.
William: Sharing ideas.
Margaret: So, I think . . . you become less isolated
because you know that you’ve got that support.

(INT: CHS, 1, FST, GRP, P. 7, L. 12–34)³

This extract illustrates two ways in which the group interview context enabled teachers to reveal the shared perspectives that they held about the purposes of moderation. At one level this was evident in the teachers’ simultaneous utterance of the word “consistency” (line 3). Similarly, a shared perspective about the educative purpose of moderation emerged as the teachers exchanged ideas about how the collaborative nature of their moderation practice had enabled them to learn from each other (lines 4–13).

³ The referencing system that I use to cite moments from transcripts of interviews (INT) is as follows. There are six parts to the reference. First the school is identified as CHS (Central Heights School), RS (Riverside School), or GS (Greenville School). Then the specific transcript that the comment is drawn from is cited. In this case the “1” denotes that it was the first interview that was conducted at this school. Next the interviewee’s participant type is identified as LT (lead teacher), FST (focus-syndicate teacher), or PL (principal). If necessary, participant types are linked using an ampersand. For example, LT&FST denotes that a lead teacher and at least one focus-syndicate teacher were interviewed together. The fourth part of the reference indicates whether the quotation has been sourced from an individual (IND) or group (GRP) interview. The final two parts of the reference indicate the page and line numbers in the transcript that the quoted excerpt has been taken from. So the reference above (INT: CHS, 1, FST, GRP, P. 7, L. 12–34) indicates that this excerpt comes from the transcript associated with the first interview that was conducted at Central Heights School. This was a group interview involving focus-syndicate teachers. The quotation cited comes from page 7 of this transcript, beginning at line 12 and ending at line 34.
Teacher questionnaire

Because it was not feasible to interview all participants, a teacher questionnaire was administered in the last term of the school year (see Appendix K). The questionnaire was conducted at this time to enable teachers to reflect on the full range of moderation activities that they had experienced. As well as eliciting background information, the questionnaire included a series of questions about participants’ experiences of and attitudes towards social moderation. These questions were designed to triangulate or test the strength of the data that were gathered during the observations and interviews. To increase confidence in any insights that this process generated, the contents of the teacher interviews and the teacher questionnaire were designed to overlap slightly, and all focus-syndicate members were asked to participate in both. In addition to seeking demographic information, the questionnaire included background questions that were intended to elicit information about participants’ prior experiences of professional learning in assessment. Both closed- and open-ended questions were included in the questionnaire. All open-ended questions required only short answers and most of the closed questions were Likert-type items. The Likert-type items utilised a scale of 1 (strongly disagree) to 5 (strongly agree).

Preparing for analysis: Transcribing observation and interview data

I chose to transcribe the audio files associated with both observations and interviews myself. I made this decision for two reasons. First, in keeping with this study’s social constructionist underpinning, I understood that these transcripts would not be “copies or representations of some original reality” (Kvale, 1996, p. 165). Because I recognised that the process of transcription would involve both choices and interpretations (Lapadat & Lindsay, 1999), I thought it was important that these decisions were consistent with my epistemological perspective. In this regard, I adopted an approach that was closer to “denaturalized” than “naturalized” transcription (Oliver, Serovich, & Mason, 2005, p. 1274). Within a naturalised approach, participants’ verbal utterances, including accents and involuntary vocalisations, are transcribed. This approach allows researchers to analyse conversational speech patterns and is associated with the idea that “language represents the real world” (p. 1274). In contrast, researchers who employ a denaturalised approach are more concerned with the meanings and perceptions that the interviewer and the interviewee(s) create and share during the interview. Therefore, those who utilise a denaturalised approach often correct participants’ grammar and
remove utterances such as stutters and pauses. Although the production of a verbatim account of participants’ conversations is a goal of the denaturalised approach, capturing the substance of these conversations is the priority.

Consistent with a denaturalised approach, I chose to remove stutters and pauses from the transcriptions that I produced. I made this decision based on feedback that I received while visiting a school to describe the study and distribute consent forms. During this meeting, I explained that I would provide each teacher with the opportunity to review and amend her or his interview transcripts. At this point, a teacher remarked upon how awkward she had found an earlier experience of this. She explained that she had felt self-conscious and embarrassed reading a transcript associated with that research because it contained all her “umms” and “ahhhs.” I did, however, decide to include what Oliver et al. (2005) have described as “nonverbal vocalizations” (p. 1283) in my transcripts. This practice, which is sometimes associated with a naturalised approach, was particularly important during the transcription of observations. During these meetings, teachers regularly referred to documents and artefacts to inform their discussions. Because I wanted to understand how teachers’ use of such items had informed their learning, I needed to have a record of the use that was made of them. To include these nonverbal vocalisations in transcripts, I listened to the relevant audio file multiple times and made repeated reference to both the field notes and the artefacts that related to that meeting. The following extract illustrates how I used non-verbal information gleaned from field notes in conjunction with school-specific artefacts to ensure that the resulting transcriptions provided a meaningful account of events:

**Kerry:** Now I gave everybody a copy of that recount [chapter], by this writer:

*Holds up a book entitled, “Write Ways.” The book is yellow with red and blue triangles on the front (Wing Jan, 2001).*

....

**Kerry:** What would we expect to see in terms of the recount [after 6 months at school]? So far, in terms of surface features, you’ve got *[reads from her school’s success criteria for personal expressive writing]* “Can record the initial or dominant sounds in words; Can record some high-frequency words” [existing
The second reason that I decided to complete the transcriptions myself was that I felt it would provide me with an opportunity to gain familiarity with the data. Although I did not develop or assign codes during the transcription process, I often wrote a brief memo after finishing each transcription (see Appendix L). In these memos, which became part of my audit trail, I recorded questions that had occurred to me while completing the transcription. I also listed areas that I felt required further exploration during analysis. In total, the transcription of observation and interview data resulted in the production of over 1000 pages of text. Each school’s transcriptions were then arranged chronologically and bound in a separate volume. Although these transcriptions were later imported into NVivo 10.2.2 (QSR International, 2014) for coding and analysis purposes, the line numbers that I had assigned to help me navigate transcriptions were lost during importation. Because of that, I continued to make frequent reference to the bound hard copies during the analysis process.

Data analysis

This section describes the methods that were used to guide data analysis and provides a rationale for their inclusion. During this phase of the research, two grounded theory practices (Charmaz, 2006) were utilised to explore observation, interview, and open-ended questionnaire data. These practices, which were initial coding and focused coding, are explained later in the chapter. Grounded theory methods, which were first developed in the 1960s (Glaser & Strauss, 1967), provide qualitative researchers with a suite of

---

The referencing system that I use to cite moments from transcripts of observations (OBS) differs slightly to the system that I employ for interview transcripts (INT). An explanation of the system for observations now follows. There are five parts to the reference. First the school is identified as CHS (Central Heights School), RS (Riverside School), or GS (Greenville School). Then the specific transcript from where the comment is drawn is cited. In this case the “5” denotes that it was the fifth moderation meeting that was observed at this school. Next the type of meeting is identified as SW (school-wide), FS (focus-syndicate), or IS (inter-syndicate). The fourth aspect of the reference indicates the page in the transcript that the excerpt comes from. The fifth and final part of the reference indicates the line numbers in the transcript that correspond to the excerpt. So the reference above (OBS: CHS, 5, SW, P. 4, L. 30–46) indicates that this excerpt comes from the transcript associated with the fifth observation that was conducted at Central Heights School. This was an observation of a school-wide meeting and the excerpt, which comes from page 4 of this transcript, was sourced from lines 30–46.
methodological strategies that can be used to generate new theory. Glaser and Strauss developed these strategies to enable researchers to inductively discover or generate theory through the identification of categories that emerge from the data. Accordingly, Glaser and Strauss maintained that researchers should resist basing even initial decisions about data collection and analysis on a “preconceived theoretical framework” (p. 45).

More recently, grounded theorists such as Charmaz (2006) have advocated using grounded theory methods more flexibly to complement other approaches to qualitative research. Charmaz’s approach, which she has described as constructivist as opposed to objectivist grounded theory, encompasses social constructionism. Her approach is premised on the understanding that researchers construct, rather than discover, grounded theories through their “past and present involvements and interactions with people, perspectives, and research practices” (p. 10). Like Wenger’s (1998) social theory of learning, Charmaz’s (2006) approach is founded on the understanding that learning is the product of social participation. Because of this, her approach to grounded theory seemed well suited to exploring how and what the groups of teachers in this study had learned through their involvement in social moderation. Grounded theory methods can be used to guide an inquiry from its inception (Charmaz, 2006; Glaser & Strauss, 1967). Within this study, however, these methods were employed to enrich the analysis process. The upcoming section will describe why and how two of Charmaz’s (2006) grounded theory practices were used to inform this phase of the analysis. Because only a small amount of quantitative data was collected during this study, frequency distribution tables were manually created to expedite the analysis of those data.

Initial coding and chronology development

During the initial coding phase, all observation and interview transcripts were comprehensively reviewed and a set of chronologies was developed (see Appendix M). Pioneered by Gunn (2008), chronology development involves the creation of “short hand versions” (p. 83) of lengthy transcripts. The development of these document road maps not only makes the subsequent identification of specific sections of text within transcripts easier but also enables the researcher to gain familiarity with the full range of topics canvassed in the data. Building on the work of Gunn, whose chronology development involved capturing the key topics and conversational turns that occurred during focus-group interviews, I used a blend of “line-by-line” and “incident to incident”
(Charmaz, 2006, pp. 50–53) coding to construct sequential, thematic outlines of each observation and interview. This process, which began while data collection was ongoing, provided me with a way of managing the mountain of transcript data that my study ultimately generated.

To create these chronologies, I worked systematically through each transcript developing a series of short, descriptive codes that emerged from the data. Because I was aware that existing notions that I had about assessment, moderation, standards, and teacher professional learning were likely to shape my responses to the data (Charmaz, 2006), I used a number of strategies to keep my approach to coding as open-ended as possible. During the construction of observation chronologies, this included using the following questions to guide my initial readings of these data. These questions, which were selected from a larger set, were:

- What is going on? What is the overall activity being studied? What specific acts comprise this activity?
- What practices, skills, stratagems, methods of operation do actors employ?
- Which theories, motives, excuses, justifications or other explanation do actors use in accounting for their participation? How do they explain to each other, not to outside investigators, what they do and why they do it? (Mitchell, 1991 as cited in Charmaz & Mitchell, 2001, p. 163)

As I read through the observation transcripts, these questions helped me to develop a series of short, active codes that deepened my understanding of the phenomenon of interest. For example, the development of codes such as “using or referring to a centrally developed resource,” “determining how to develop an assessment task,” and “referring to relevant professional development experiences” helped me to identify the practices and skills that participants employed as they engaged in moderation activity (second bullet point). This contributed to my understanding of the overall suite of activities that were associated with moderation practice (first bullet point). Similarly, the creation of codes such as “identifying a purpose or rationale for moderation” and “valuing opportunities to collaborate” provided some initial insights into the theories and justifications that participants used to account for their participation in moderation activities (third bullet point). Because codes that capture participants’ actions help to preserve their experiences (Charmaz, 2006), I used verbs and gerunds to keep these initial codes active across both observation and interview chronologies. Within each
chronology, the judicious inclusion of representative quotes from the data served not only to define these codes but also to ensure that they were anchored around the participants’ meanings and actions.

Throughout the process of chronology development, I employed constant comparative methods (Glaser & Strauss, 1967). This involved comparing data with data both within and across cases. By constantly reviewing and revising chronologies as I worked through each observation and interview transcript, I ensured that these initial codes were used consistently. During this process, I also created a separate “initial codes” document (see Appendix N). Within this document, which evolved as each chronology was developed, I grouped the initial codes according to the type of activity, process, or perception they described. I assigned each of these groups a brief title, or category code. For example, the category code “involving students” was assigned to the following initial codes: employing “kid speak,” sharing criteria with students, promoting self-assessment, developing students’ assessment capability, and promoting sharing student writing with other classes. To conclude the initial coding phase, I reviewed all responses to the open-ended questionnaire items to identify whether any new codes needed to be added to the initial codes document. Because no further codes were identified during this phase, I then began a process of focused coding.

Focused coding

Charmaz (2006) explained that focused coding involves selecting the most salient initial codes and using these to review and code large amounts of data. During this process, I used the chronologies, often in combination with the initial codes document, to explore what the open-ended questionnaire responses and the observation and interview transcripts revealed about learning as a process of doing, experiencing, becoming, and belonging (Wenger, 1998). This involved reading the data in four related ways: to identify the practices that had sustained teachers’ mutual engagement in moderation; for instances in which participation in moderation had enabled teachers to negotiate meaning about assessment; to explore how taking part in moderation had shaped teachers’ identities as assessors; and for insights into whether – and if so, how – teachers’ involvement in moderation activities had contributed to the development of assessment-focused communities of practice.

Because I understood that factors such as the provision of extended time for opportunities to learn, the involvement of an external expert, the existence of active school leadership, and the provision of a rationale for engagement in learning
(Timperley et al., 2007) might affect how and what teachers learned through their involvement in moderation, I also read the data looking for examples of these factors. During each reading of the data, I drew upon the concepts of dependability, reliability, and validity, as they were defined in Chapter 2. By using these definitions in conjunction with understandings about how standards are promulgated (Sadler, 1987), I was able to examine the data for evidence of specific ways in which teachers had used their involvement in moderation to learn about assessment. As I reviewed the data, the influence of teachers’ understandings of assessment for learning principles emerged as an important theme. Further details about how this theme was both identified and accommodated for are provided later in this section.

Before embarking on the focused coding phase, I imported all open-ended questionnaire responses, transcriptions, and chronologies into a computer programme called NVivo 10.2.2 (QSR International, 2014). Using this programme, which is designed to support and expedite qualitative data analysis, I iteratively reviewed the imported data and assigned descriptive codes to the text. To do this, I selected from the inductively derived codes that were identified during the initial coding phase. Each time a new code was assigned, all previously coded transcripts and questionnaire responses were reviewed to capture incidences that matched the introduced code. Throughout the focused coding phase, I used the chronologies to help me navigate observation and interview transcripts. For example, in one early reading of the data I noticed that Kerry spoke enthusiastically about the benefits associated with using involvement in moderation to develop local assessment criteria. She said,

One of the benefits of what we are doing is that the success criteria are very home based. It’s to do with the genre that we are writing, the teaching that we’re putting in, and it helps shape the teaching that we do put in there.

(OBS: CHS, 8, SW, P. 21, L. 16–18)

Because I wondered if Kerry’s statement might provide an insight into what had sustained other teachers’ mutual engagement in moderation, I turned to the chronologies to help me identify other similar instances in the data.

To do this, I reviewed the chronology that was associated with this statement and noted that I had previously associated it with two codes. These were the category code “strengthening teaching” and the initial code “valuing school-specific processes or products.” Next, using the “query” function that NVivo 10.2.2 (QSR International, 2014) offers, I was able to search all my chronologies for other instances of these codes.
By utilising the rigorous referencing system that I had employed within my chronologies, I quickly located each of these instances in the observation and transcript data. I then systematically reviewed each section of text and assigned the appropriate code or codes to it. Using a different feature of NVivo 10.2.2, I then extracted all the text passages associated with these codes and generated a node report for each. The following are examples of the kinds of passages that were assigned to each of these codes:

**Margaret:** My assessment practice has improved since we have been focusing on this [moderation]. . . So, it’s [moderation] got to be good. If my assessment practice is improving, then my teaching is going to improve from that. <stressing teaching>
(INT: CHS, 1, FST, GRP, P. 38, L. 42–51)

**Susan:** Two kids today got up and went to have a look at that [set of criteria developed during a moderation session]. And I’ve only just introduced that [those criteria] . . . They were doing their recount and they got up and went and had a look at that and came back. So that was good. <valuing school-specific processes or products>
(OBS: GS, 2, IS, P. 36, L. 37–46)

**Kerry:** I think it’s really good that we share those [success criteria] with the kids when they are doing their writing. Start working towards that. <valuing school-specific processes or products>
(OBS: CHS, 8, SW, P. 21, L. 14–16)

**Chris:** When you have got those [success] criteria . . . you realise . . . that’s what they [the students] are supposed to achieve. . . . You can work out what [the] little steps [are], like what’s your next [teaching] step? <valuing school-specific processes or products> <stressing teaching>
(INT: CHS, 2, FST, IND, P. 12, L. 40–43)

As I looked across these reports, I noticed that I had assigned both of these codes to a number of statements. This made me wonder whether there was a common idea that linked or underpinned these two codes.

To explore this, I examined how the statements related to each other. For example, I noted that although Kerry (first statement), Margaret, and Chris all identified ways in which involvement in moderation had positively informed their teaching
practice, Kerry and Chris talked about the crucial role that their local success criteria had played in this regard. Likewise, Susan spoke about a benefit associated with the assessment criteria that she and her colleagues had developed during their involvement in moderation. Her interest, however, was in the way in which the students in her class had used these criteria to take increased responsibility for their learning. This was an idea that Kerry alluded to in her second statement. Here, she reminded her colleagues to share their local assessment criteria with students so that they could use these to inform their writing.

As I reviewed these two node reports, I realised that all the statements related in some way to the principles of assessment for learning (Broadfoot et al., 2002). That is, they either showed how teachers had used involvement in moderation to inquire into and inform their teaching, or revealed how moderation had encouraged or helped teachers to communicate learning goals to their students. To explore this further, I checked the initial codes document and noted that a number of other codes appeared to be aligned with these principles. These included the initial codes associated with the category “involving students” and a series of initial codes that related to the identification of next steps. Pursuing this lead, I used the procedure that was outlined earlier to systematically identify and code instances in which teachers either involved students or identified next steps. During this review of the data, the codes “involving students” and “identifying next steps” were assigned to these instances. Having completed this systematic review, I generated two further node reports and examined these alongside the two reports that I had produced earlier. On this basis, I created a parent node entitled “assessment for learning” to encapsulate the “strengthening teaching,” “valuing school-specific processes or products,” “involving students,” and “identifying next steps” nodes. The “assessment for learning” theme was then brought into my readings of the data and all previously coded transcripts were reviewed accordingly. Although this process was replicated on many occasions, it became apparent that the assessment for learning theme was particularly important. For this reason, I subsequently re-examined the research on moderation and added a section about moderation and assessment for learning into my review of the literature.

**Ethical considerations**

Before undertaking this research, the project was submitted for consultation with the University of Otago Ngāi Tahu Research Consultation Committee. At this time, approval for the project was also sought and gained from the University of Otago.
Human Ethics Committee. In accordance with this institution’s ethical guidelines, I provided information sheets to all participants and later collected signed consent forms from each person. During this process, consent was gained from the principal as well as from all teachers at each school. Additionally, each school’s Board of Trustees was provided with an information sheet and written consent was obtained from the board chairperson at each school. All participants were aware that their participation in the study was voluntary. Similarly, they understood that they were free to withdraw from the project at any time. Because I observed teachers discussing student writing and anticipated that I might need to collect anonymised samples of student work, information about the study was also provided to the parents and guardians of the children at each school. This sheet explained that parents and guardians could request for their child’s/children’s work not to be involved in the study. It also made it clear that such a request would not disadvantage their child/children in any way. As noted earlier, samples of all the information sheets and consent forms that were associated with this study are included in Appendix A.

Participating in qualitative research can carry unanticipated risks for participants (Merriam, 2009). For example, participants sometimes disclose more than they mean to during interviews. To address this, each person who took part in an interview was provided with the opportunity to review and amend the resulting transcription prior to data analysis. Furthermore, all participants were provided with the contact details for my primary supervisor and myself so that they could discuss any concerns that might arise during their involvement in the study. Although it is never possible to guarantee anonymity, steps were taken to protect the privacy of both the schools and the participants. These included presenting findings as specific cross-case issues rather than through individual case reports. It was hoped that this would decrease the likelihood that an individual school would be identified. For the same reason, pseudonyms have been used throughout this thesis for the names of all of the schools and participants.

Study and researcher effect

Research conducted from a social constructionist perspective should address issues that arise from the interdependent relationship that exists between the researcher and the research participants (Marechal, 2010). This is especially true within case study research, in which the researcher becomes a part of the process and the product of the research (Begoray & Banister, 2010). Case study researchers who adopt a reflexive approach systematically document their interactions with participants and reflect upon
how their participation may have shaped the research outcomes. They also seek to
acknowledge the ways in which their own biases and assumptions could have affected
both the process and the product of the research. Within this section I reflect upon how
my participation, and even the idea of being involved in my study, may have affected
the moderation processes that were enacted at each of the participating schools. This
section also explains how these effects were accounted for during data analysis. Three
main effects are examined: the study’s influence on each school’s uptake of moderation
activity within the National Standards context, the possible effect that my presence had
on the leadership of moderation processes, and the impact that my participation may
have had on the understandings that the teachers at one school arrived at in relation to
the selection and use of standardised assessment tasks. To conclude this section, I
describe a further measure that was taken to test the credibility of my interpretations of
the data.

As outlined in the previous chapter, the 2010 introduction of the National
Standards (Ministry of Education, 2009c) set the scene for this study. Although many
New Zealand primary school teachers had some prior experiences of engaging in
moderation (Wylie & Hodgen, 2010), the introduction of the National Standards
focused the sector’s attention on this assessment process. Within this new assessment
landscape, there were heightened expectations that teachers would engage in social
moderation activities (Ministry of Education, 2010e). For these reasons, the
introduction of the National Standards provided an opportune context for my research.
In planning this study, I sought to work with schools that were actively engaging in
moderation activities. Because several studies examining the early implementation of
the National Standards had identified that teachers at most primary schools were
engaging in the moderation of student writing by the end of the 2010 school year
(Education Review Office, 2010; Thomas & Ward, 2011), I expected that this would
also be true of the participating schools.

Although the principal of each participating school indicated that teachers would
be moderating student writing during the 2012 school year (the data-collection phase), it
became apparent that this represented each school’s first attempt to undertake
moderation activity since the introduction of the National Standards. Given this, it is
hard to determine the extent to which involvement in my study prompted the
participating schools to focus their attention on social moderation at this time. Here, the
question is not whether but when these schools would have updated their moderation
processes had they not been involved in this study. Certainly, one teacher at Greenville
School commented, “If you hadn’t have been coming, I don’t know if we would have got going so quickly [with our moderation work]” (INT: GS, 1, FST, GRP, P. 54, L. 24–25). Also hinted at in the excerpt above is the related question of whether my study affected the frequency of each school’s moderation activity. This would be hard to determine. There was no indication that schools only engaged in moderation activity when I was present. Although I made every effort to attend all of the moderation meetings that I was informed about, I did miss at least one meeting at each school. It is, however, unclear whether schools planned and undertook more moderation meetings than they would have if they had not been participating in the study.

The second point relating to the possible impact of my study on each school’s uptake of moderation activity concerns the selection of lead teachers. As described earlier, I asked the principal at each school to identify the teacher with existing responsibility (independent of this study) for organising and leading moderation sessions. This teacher was invited to become the lead teacher for the study. My assumption was that these teachers would have been leading their school’s moderation processes for some time. Yet because this study’s data-collection phase coincided with each school’s first attempt to engage with moderation within the National Standards context, this was not the case. Although I understood that the teachers who assumed this position did so because they had expressed an interest in moderation prior to the first contact that I made with their school, it is possible that this aspect of my study may have shaped how moderation processes were led.

To address this, I took extra care during the analysis phase to identify the ways in which participants negotiated and enacted their leadership roles. This raised questions about the effect that my study may have had on the way in which moderation processes were led at one school, Greenville School. In particular, I wondered whether Susan, who became a lead teacher for this study after independently expressing an interest in developing her school’s moderation processes, would have gone on to lead these activities if the study had not taken place. Regardless of whether this study had an unintended effect on how moderation processes were led at Greenville School, this school’s approach rendered some important insights into how leadership issues can shape the opportunities that moderation affords teachers to learn about assessment. This is dealt with in greater detail in Chapters 5 and 7.

Begoray and Banister (2010) noted that reflexive case study researchers should “attend to an ethic of care for those who participate in the research” (p. 788). Yet to be reflexive in an ethical sense, researchers must identify and respond to the ethical
tensions that occur (Guillemin & Gillam, 2004). The final point relating to the possible impact of my study concerns an ethical tension that arose at Riverside School. Specifically, it stems from an event that occurred during the final syndicate-level moderation meeting that I observed at this school (see Appendix G, Meeting 3). During this meeting, a group of teachers developed a process that they believed would allow them to convert the scores that their students had attained on a series of three NEMP tasks (Crooks et al., 2007) into a National Standards-type rating (e.g., well below, below, at, and above). At this time, the teachers were preparing written achievement reports to send home to parents and guardians. This is a process that New Zealand primary school teachers are expected to undertake twice each year (Ministry of Education, 2009d). My understanding was that the National Standards-type rating, which had been generated from each child’s scores on the NEMP assessments, would be included in his or her individualised report of achievement. As this syndicate meeting drew to a close, Stephen, who was this school’s lead teacher, noted that it would be good to meet as a full staff to discuss the way in which they had used the NEMP tasks. I was invited to attend this full-staff meeting.

At the time of data collection, I was also working as a researcher for the organisation that had developed the assessment tasks that the Riverside School teachers were using. Although I was acutely aware that the teachers’ use of the data associated with these assessment tasks was not appropriate, I did not comment upon this during the meeting that has just been described. Yet as I later reflected upon the teachers’ use of the NEMP data, I realised that I felt ethically bound to provide them with some feedback about this. Given that these teachers knew about my association with the organisation that had produced the NEMP tasks, I was concerned that they might interpret my silence as an assurance that their interpretation and use of these data was valid. In consultation with my primary supervisor, I emailed Stephen and outlined the concerns that I had about their use of the NEMP data. Within this email, I described an approach that would enable the teachers to use the NEMP data to arrive at valid National Standards-type judgements of student achievement. I also offered to talk to other teachers about this approach at the upcoming full-staff meeting. In his reply, Stephen indicated that he would like me to come to the meeting. He also asked me to talk to his colleagues about how they might use the approach that I had outlined to inform their subsequent use of NEMP assessment tasks. I agreed to do this.

I prepared carefully for that meeting. Because I wanted to ensure that I had a full record of the information that I shared with the teachers during this session (see
Appendix G, Meeting 4), I took even more care than usual to ensure that my audio-recording devices were functioning as intended. This full-staff meeting allowed the principal, who had planned the moderation process under discussion, to clarify how she had intended the NEMP tasks to be used. Although Stephen had independently scheduled this meeting to discuss his syndicate’s use of these NEMP tasks, it is unclear whether my contribution to it influenced the principal’s response. Likewise, it is possible that my participation may have shaped an understanding that Kim, one of this school’s focus-syndicate teachers, arrived at regarding the selection of assessment tasks. Further details about the learning opportunities that involvement in these events may have afforded the participating teachers are presented in Chapter 5.

There were a number of other indications that my presence may have had an impact upon teachers’ experiences of moderation. For example, a teacher from Central Heights School commented that she thought my presence had motivated teachers to reflect more during a judgement-making session. Similarly, another teacher from the same school commented that student writing had been read aloud during all judgement-making sessions for my benefit. As noted earlier, I was at times asked to answer questions during moderation sessions. Yet all of the questions that I responded to were of a very minor nature and did not prompt changes in teachers’ moderation practice. Although it is difficult to measure how these factors, such as being motivated both to read writing aloud and to reflect more upon judgements, shaped teachers’ experiences of moderation, it seems unlikely that they had a major effect on the opportunities that participating in these processes afforded teachers for learning about assessment.

Once data collection had been completed, I had the opportunity to lead feedback sessions at both Central Heights and Greenville schools. As well as helping to ensure that participants benefited from their involvement in the study, the feedback sessions provided me with an opportunity to share my understandings with participants. Because I knew that these understandings were based upon my interpretations of their experiences of social moderation, I recognised that these feedback sessions also served a member-checking function. Merriam (2009) explained that researchers often employ member checks, which involve seeking feedback from participants about preliminary analyses, to ensure the credibility of their findings. Because of the participative nature

5 Although I had been scheduled to lead a session at Riverside School, this was postponed due to my ill health. Unfortunately, it was not possible to reschedule this session. As noted earlier, I had previously provided the teachers at Riverside School with preliminary feedback.
of the feedback sessions that I facilitated, conducting these sessions had the effect of increasing the confidence that I had in my findings.

**Chapter conclusion**

This chapter has introduced and explained the various methods that were used to produce and analyse data within this qualitative study. These processes allowed me to explore how and what the teachers at three New Zealand primary schools learned about assessment through their involvement in social moderation. The findings from this study, which are presented in the following three chapters, suggest that participation in social moderation can provide teachers with the opportunity to strengthen their assessment capability. Yet these findings also highlight the complexity of moderation practice and reveal some of the learning barriers that teachers encountered while undertaking this work.
Chapter 4. Moderation as assessment practice: Using involvement in moderation to build teachers’ assessment for learning capability

This chapter draws on Wenger’s (1998) conception of practice, or “learning as doing” (p. 5), to examine how and what teachers learned about assessment for learning principles and practices through their involvement in social moderation. It explores how the resources, systems, and perspectives that teachers developed and used to sustain their mutual engagement in moderation shaped the learning opportunities that they were afforded. The chapter opens by describing the learning benefits associated with using involvement in moderation for conceptualising and developing assessment resources. First, it examines how the teachers at Greenville School used their conversations about assessment criteria to develop a system for fostering students’ assessment capability. Next, it describes how the teachers at Central Heights School used their involvement in resource development to strengthen their ability to construct success criteria. The chapter concludes by explaining how taking part in moderation enabled the teachers at Central Heights School to get better at using judgement-making sessions for identifying next teaching and learning steps. These findings, which are drawn from the two participating schools that aligned their moderation processes with assessment for learning principles and practices, suggest that involvement in moderation can provide teachers with assessment-focused professional learning opportunities. On this basis, it is argued that engaging in social moderation can strengthen teachers’ assessment for learning capability.

---

To avoid ambiguity, the term *system* is used here in place of the word *framework* (see Wenger, 1998, p. 5). The term framework has multiple applications within the moderation literature. For example, James (1998) has presented a multi-stage framework for school-based moderation. This framework is procedural and proposes activities that teachers might undertake before, during, and after assessment. In contrast with this, Maxwell (2002b) has used the term framework to refer specifically to the standards or criteria that student work is appraised against. Within this thesis, the standards and criteria that teachers use to inform their judgements of student work are referred to as *resources or tools*. Likewise, the procedures and approaches that teachers employ during their moderation practice are described as *systems*. 
Conceptualising a moderation resource: Developing a system for building students’ assessment capability

At Greenville School, the process of conceptualising and constructing a moderation resource enabled teachers in the middle (years 2–4) and senior (years 4–6) syndicates to focus their attention on the development of students’ assessment capability. Specifically, these teachers used their involvement in moderation to develop a system for encouraging students to take increased responsibility for monitoring their own learning. These self-monitoring skills are important because students need to be able to inquire into and assess their own learning to become self-regulating, autonomous learners (Broadfoot et al., 2002; Sadler, 1989). The development of students’ assessment capability has been a priority for some time in New Zealand. For example, many of the teachers who participated in the Ministry of Education’s long running Assess to Learn (AToL) project used their involvement in this initiative to learn how to encourage students to use self- and peer-assessment (Poskitt & Taylor, 2008). Likewise, the publication of the Directions for Assessment in New Zealand report (Absolum et al., 2009) drew attention to the importance of building students’ assessment capability (Ministry of Education, 2011c). The findings that are presented next show how the teachers at Greenville School used their participation in moderation to strengthen their ability to involve students in the assessment process.

Although the teachers at this school reported that they had not taken part in the Ministry of Education’s AToL initiative, it was evident that they had some familiarity with the principles and practices of formative assessment. For example, during a number of their moderation meetings (see Appendix F, Meetings 2 & 6), teachers in the middle (focus) and senior syndicates discussed ways to integrate their existing use of learning intentions into the moderation system that they were developing. Notably, many of the teachers at those schools that did participate in the AToL project used their involvement to learn how to develop and co-construct learning intentions with their students (Poskitt & Taylor, 2008). Similarly, the teachers at Greenville School also made reference to their use of “I Can” charts. These charts, which had been developed prior to the introduction of the National Standards (Ministry of Education, 2009c), were written in child-friendly language and listed the writing goals that children were working towards. Some of the teachers at this school attached these charts to the

---

As explained in Chapter 3, I use the term focus syndicate to refer to the teaching team that I worked the most closely with at each of the participating schools.
children’s writing books and highlighted or ticked each goal as it was achieved. The teachers’ use of approaches such as these provided an indication of their familiarity with the principles of formative assessment. The rest of this section describes how these teachers were able to use their involvement in moderation to strengthen their knowledge and understanding of these principles and practices.

In this regard, the teachers at Greenville School benefited enormously from the opportunities that participating in moderation afforded them for working alongside a colleague who had extensive assessment for learning expertise. This colleague, Phillip, explained that he had gleaned much of his assessment knowledge while working as a teacher in the United Kingdom. During this time, he had taken part in a major professional development contract focused on assessment for learning. Through his participation in this initiative, Phillip was exposed to a variety of strategies for fostering both self- and peer-assessment. While participating in Greenville School’s moderation sessions, he shared a number of these strategies with his New Zealand colleagues. According to Wenger (1998), communities of practice act as locally defined “regimes of competence” (p. 136). Within such communities, notions of competence are established and negotiated through each group’s practice. Because Phillip was able to use moderation meetings as a forum for introducing his colleagues to new assessment for learning practices, he helped them to redefine and expand their regime of competence. In this way, Phillip’s expertise was one of the resources that these teachers utilised to sustain their mutual engagement in moderation.

In particular, the teachers at this school drew on Phillip’s expertise to inform the approach that they developed to guide their moderation practice. This resulted in the construction of a series of resources that these teachers referred to as Must, Should, Could charts (see Appendix C). In keeping with the advice that the Ministry of Education had provided to schools at this time (Ministry of Education, 2011a), the teachers at Greenville School used the New Zealand Curriculum Reading and Writing Standards (Ministry of Education, 2009c) and the Literacy Learning Progressions (Ministry of Education, 2010b; henceforth referred to collectively as the centrally developed writing standards) to inform the construction of their Must, Should, Could charts. Because of this, each of these resources became part of the “shared repertoire” that the teachers at this school used to inform their moderation practice (Wenger, 1998, p. 73). Of these resources, Greenville School’s locally developed Must, Should, Could charts were of particular importance. These charts, which were written in child-friendly language, described what student writing must, should, and could exhibit by the end of
a given year level at Greenville School. Within this system, those learning goals listed in each chart’s could category described those skills and abilities that were required at the next year level. Further details about the development of the criteria that were specified on this school’s Must, Should, Could charts, including the ways in which the teachers sought to align these charts with the centrally developed writing standards (Ministry of Education, 2009c, 2010b), are provided in the following chapter.

The Must, Should, Could approach was introduced to teachers during Greenville School’s initial moderation meeting (see Appendix F, Meeting 1). In this session, teachers from the middle and senior syndicates discussed the resources that they would develop to inform their judgements of student writing. As was the case with all the moderation meetings that were observed at this school, Susan, who was also the leader of the middle syndicate, took responsibility for facilitating this session. Early in the meeting, Susan prompted Phillip to explain an assessment approach that he had previously described to her:

1 Susan: And I liked what Phillip was saying . . . too. It’s really nice to share ideas. . . . You know how you sort of put down in your criteria for the kids, and you had “must have.”
2 Phillip: Oh the musts, shoulds, and coulds . . . Everybody “must” do this, some of you “should” be able to do that and some of you “could” go on to do that.
3 Susan: And I quite liked that. I think the kids would get that. . . . And I think it sets that, trying to get up to the next level for the kids. So I think that’s the way to go. . . . So will we get together early next term and look at [developing]
4 some?
5 Jack: Yeah, that would be good.

(OBS: GS, 1, IS, P. 29, L. 8–32)

Here, Susan played an important role in creating an opportunity (lines 1–3) for Phillip to use his involvement in moderation to share an assessment idea (lines 4–6). Because Susan had previously identified that Phillip’s Must, Should, Could system had the potential to strengthen students’ assessment capability (lines 7–9), she also used this meeting as an opportunity to canvass her colleagues’ interest in this approach (lines 9–11) and seek their buy-in (line 12). This illustrates the way in which moderation meetings can provide teachers with a forum for sharing assessment ideas and expertise.
Teachers in the middle and senior syndicates met again later to further refine their interpretation of this approach (see Appendix F, Meeting 2). During this meeting, Phillip drew on his prior experiences of using the Must, Should, Could system to alert his colleagues to the ways in which this approach could be used to build students’ assessment capability. Referring to his experiences in the United Kingdom he explained, “When you get them [the students] to buy into that [the Must, Should, Could system], it’s brilliant because then they start being self-evaluative. And that’s when they really take off” (OBS: GS, 2, IS, P. 2, L. 23–24). He later elaborated and clarified that this was possible because, “the kids have got such a clear framework that they can look at [with the Must, Should, Could charts]. You can get them to self-assess, you can get them to peer-assess” (OBS: GS, 2, IS, P. 39, L. 25–26).

Evidence-based research on teacher professional learning and development has indicated that teachers require a catalyst or a rationale to motivate them to engage meaningfully in professional learning opportunities (Timperley et al., 2007). At Greenville School, Phillip’s ability to connect participation in moderation with opportunities for teachers to build students’ assessment capability seemed to provide his colleagues with the rationale they required to sustain their involvement in moderation activity. Certainly, both of Phillip’s middle-syndicate colleagues seemed to readily identify that the Must, Should, Could approach was one that their students would buy into. Prior to formalising their syndicate’s Must, Should, Could charts, both Susan and Sophia trialled this concept with their classes. Here, Sophia, a teacher of year 2 and 3 students, described her class’s initial response to the Must, Should, Could approach:

1  I put “could underline the tricky part of a word” just because
2  I’d just introduced it. . . . I wasn’t expecting any of them to
3  try it. Every single one of them, except for two, tried
4  underlining the tricky part of the word, because it was on the
5  “could”. . . . And that was amazing, like normally I’d have
6  to drum that in and make a thing about it, you know. They
7  all tried it.

(OBS: GS, 2, IS, P. 38, L. 25 – P. 39, L. 10)

For Sophia, the realisation that this approach had motivated most of her students to engage with assessment criteria (lines 3–4 & 6–7) seemed to provide her with an incentive or a rationale (lines 5–7) for continuing to engage in moderation practice.
This was a perspective about the Must, Should, Could system that others appeared to share. Almost a month later, when Susan arranged to use part of a full-staff meeting to promote their moderation approach to teachers in the junior syndicate (see Appendix F, Meeting 5), she and her colleagues emphasised the motivational aspect of the Must, Should, Could system. By this stage, the teachers in the middle and senior syndicates had used their involvement in moderation to develop a series of Must, Should, Could charts for students in years 2–6. During this full-staff meeting, Susan described their evolving system and told her colleagues about how her students had responded to it. She explained, “I’ve had kids who have gone and looked at this [the Must, Should, Could charts], just automatically on their own, just to see if they were doing the right thing” (OBS: GS, 5, SW, P. 2, L. 7–8). For Susan, the idea that the Must, Should, Could charts had encouraged her students to take increased responsibility for evaluating their own learning was very important. Likewise, Jack, a teacher of year 5 and 6 students, was also impressed with the way in which this system seemed to have motivated his students to engage with their assessment criteria. He explained to his colleagues, “It’s a good motivator, the ‘should’ and ‘could’. Cause . . . the children want to be out there. They want to be doing stuff that is a bit special” (OBS: GS, 5, SW, P. 5, L. 9–10).

The data that have been presented within this section demonstrate that teachers in the middle and senior syndicates at Greenville School shared the view that the Must, Should, Could approach had encouraged their students to take increased responsibility for monitoring their own learning. For these teachers, the shared perspective that their involvement in moderation was motivating students to take increased responsibility for their learning seemed to provide the teachers with the rationale they required to sustain their engagement in this assessment activity. The development of this shared perspective also provides an example of what Wenger (1998) would describe as the “negotiation of a joint enterprise” (p. 77). According to Wenger, a community’s joint enterprise is “defined by the participants in the very process of pursuing it” (p. 77). This was certainly the case at Greenville School, where the joint enterprise that the teachers in the middle and senior syndicates negotiated was inextricably linked with their involvement in the development of their moderation approach. Because of the expertise that Phillip shared with his colleagues, these teachers were able to use their involvement in these processes to develop a new system for fostering students’ assessment capability.
Constructing assessment resources: Strengthening teachers’ criteria-development skills

This section describes how involvement in moderation enabled teachers at Central Heights School to strengthen their ability to develop assessment criteria. Like their colleagues at Greenville School, the teachers at Central Heights School used a number of their moderation meetings for developing and reviewing local assessment resources. At this school, these meetings (see Appendix E, Meetings 1, 2, 5, & 12) were used to construct rubrics that detailed the criteria that teachers then used to inform the teaching and assessment of writing. The teachers at Central Heights School referred to these rubrics as their Success Criteria (see Appendix B). The New Zealand Ministry of Education has explained that success criteria should provide students with a clear understanding of the qualities that teachers will look for when making judgements about their work (Ministry of Education, n.d.-h). To help students establish these understandings, teachers are encouraged not only to share success criteria with their students but also to involve them in the development of these criteria (Ministry of Education, n.d.-h). This approach is consistent with Sadler’s (1989) theory that a student can only achieve a learning goal if she or he understands that goal and is able to assess what is required in order to achieve it.

For students to be successfully involved in the development of learning goals, such as success criteria, teachers first require a clear understanding of the role and function of these descriptions of achievement. Similarly, teachers need to be aware of those properties that are associated with effective success criteria. Many of the teachers who took part in the Ministry of Education’s AToL project used their involvement in this initiative to learn how to develop and co-construct learning intentions and success criteria with students (Poskitt & Taylor, 2008). Although Central Heights School participated in this initiative, their involvement was at least 7 years prior to this study’s data-collection phase. The findings reported here show how involvement in moderation focused teachers’ attention on success criteria and expanded their understanding of the function and properties of these descriptions of achievement.

At Central Heights School, the teachers developed a system of using the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to ensure that their local criteria were aligned with nationally prescribed expectations. Kerry, who as this school’s deputy principal also led their moderation processes, prompted her senior-syndicate colleagues to establish this system during their first moderation
meeting (see Appendix E, Meeting 1). Within this meeting, the teachers in the senior (focus) syndicate worked collaboratively to draft their first set of Success Criteria. During this process, these teachers were observed making 47 references to the centrally developed writing standards. Wenger (1998) has explained that a community’s shared repertoire is comprised of both the tools and the actions that it develops or adopts to inform its practice. Because the teachers in the senior syndicate used the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to systematically inform the co-construction of their first set of Success Criteria, these tools and actions became part of their emerging shared repertoire.

Having established this repertoire, these teachers then helped to ensure that all their colleagues utilised it. An example of this can be seen in the following exchange, which occurred during the first criteria-development session that involved the full staff (see Appendix E, Meeting 2). Here, Margaret, a member of the senior syndicate, used the criteria-development system that she and her colleagues had established to frame an inquiry about a criterion:

1. **Carl:** [Reading from draft Success Criteria for year 3]
2. Accurate use of capital letters, full stops, question marks, exclamation marks, some use of speech marks.
3. **Margaret:** It was actually the speech marks I was wondering about. . . . [In] the [Literacy Learning] Progressions it says after 3 [years]. . . . “Using capital letters, full stops, question marks, and exclamation marks correctly.”
4. **Carl:** Yeah, but we bang on about speech in the middle syndicate and we do teach where to put the speech marks.
5. **Kerry:** That [year 2 and 3] might be the teaching period. . . .
6. But if they [the students] couldn’t get it [speech marks] . . . that’s not going to be achieving. So, if we could just leave that one out and put it [speech marks] in when . . . they’ll independently put that in.

(OBS: CHS, 2, SW, P. 16, L. 25 – P. 17, L. 3)

As well as foregrounding the role that teachers in the senior syndicate played in extending the use of this criteria-development system, this conversation also provides an indication of the benefits associated with using the centrally developed writing standards as a basis for these conversations.
Within this exchange, it was Margaret’s reference to a centrally developed resource (lines 5–7) that prompted Carl to reveal the personally held expectation that had informed the development of his syndicate’s criterion (lines 8–9). This is in keeping with Harlen’s (1994) suggestion that the educative function of moderation is greatest when teachers have opportunities to examine the conceptions and assumptions that inform their judgement making. Although no judgements of student work were being made at this time, involvement in this conversation enabled Carl to explain his reasons for proposing a judgement-making rationale. This revealed a mismatch between his personally held expectations about achievement and those that were specified in the centrally developed writing standards (Ministry of Education, 2010b). As well as enabling Carl and his colleagues to align their expectations with those stipulated in the centrally developed documents, this experience provided all the teachers with the opportunity to reflect on and strengthen their understandings of the purpose of success criteria (lines 10–14). Specifically, this conversation prompted Kerry to remind her colleagues that these criteria are supposed to describe the qualities that teachers will look for when arriving at judgements of student work. This understanding of the role and function of success criteria is consistent with the way in which the Ministry of Education has described their purpose (Ministry of Education, n.d.-h). In this way, Kerry’s statement both clarified the role of success criteria and justified making use of the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to inform these local resources.

Although most of the teachers at Central Heights School were observed contributing to the development of their local assessment criteria, Kerry played a particularly important role in this process. Described by this school’s principal as their assessment “guru” (INT: CHS, 4, PL, IND, P. 8, L. 13), Kerry had previously participated in multiple assessment-focused professional learning opportunities. Indeed, Kerry noted that she had drawn on these experiences to inform her approach to criteria development. On many occasions, it was the guidance that Kerry provided and the questions that she asked that led to these processes providing teachers with opportunities to hone their criteria-development skills. In this respect, Kerry’s skill set was one of the resources that shaped how and what the teachers at this school learned about assessment through their involvement in moderation. For example, during the senior syndicate’s first criteria-development session, Kerry repeatedly encouraged her colleagues to tighten the specificity of the criteria that they were proposing. This approach is evident in the following exchange, which occurred after Margaret used the
Literacy Learning Progressions (Ministry of Education, 2010b) to propose a year 4 criterion:

1. **Margaret:** Use simple and compound [sentences].
2. **Kerry:** Accurate. Can write. Let’s be really distinctive, ’cause we want to be able to say “yep,” they’ve met that criteria.
3. **Accurate simple and compound sentences.**
   (OBS: CHS, 1, FS, P. 7, L. 26–31)

In her response, Kerry emphasised the importance of crafting criteria that were both clear (line 2) and measureable (line 3). In this way, she not only drew her colleagues’ attention to the features of effective success criteria but also provided them with a practical reason (line 3) for composing criteria that possessed these qualities.

Despite the fact that Kerry often took responsibility for alerting her colleagues to those features of their criteria that needed strengthening, she did not always supply the solution. As can be seen in the following exchange, this gave others the opportunity to refine their criteria-development skills.

1. **Carl:** And the last one [success criterion] was “can plan their writing and use a process of revision, editing and proofreading.” . . .
2. **Kerry:** What if they can do one of those? They’re quite different things. If you’re assessing them and they can do one, they can plan their writing, what are you going to give them? Achieved or not achieved? . . .
3. **Jean:** So the planning one and the editing one.
4. **Mia:** Make them two different, yeah.
5. **Carl:** All right.
6. **Kerry:** ’Cause they are quite different.
7. **Mia:** Both of them are equally important.
   (OBS: CHS, 2, SW, P. 18, L. 21 – P. 19, L. 5)

During this conversation, the speed with which Kerry accurately identified that the proposed criterion sought to measure two distinct skills (lines 4–5) provides further evidence of her criteria-development expertise. Rather than suggesting a solution, Kerry’s response alerted her colleagues to a challenge that the existing criterion could present (lines 4–7). This enabled Jean and Mia to cooperatively identify a solution (lines 8–12).
Because the conversation evolved in this way, opportunities were created for multiple teachers to think critically about the features of an effective success criterion.

The examples that have been presented up to this point demonstrate how involvement in the construction of school-specific assessment criteria enabled teachers to strengthen their criteria-development skills. On at least one occasion, however, engaging in these processes had an additional, unintended outcome. In this instance, the process of tightening a success criterion inadvertently resulted in the creation of an incorrectly levelled description of achievement. This occurred during a full-staff criteria-development session as teachers reviewed and revised their Success Criteria for recount and report writing (see Appendix E, Meeting 5). As was the case in the previous criteria-development cycle, teachers had met earlier in their syndicates to draft these criteria. In the exchange that follows, Carl sought to renegotiate the year 2 spelling criterion. He did this after noting the way in which the year 4 spelling criterion had been specified:

1. **Kerry:** “Correctly spells high-frequency words from the essential lists 1–4” [reading from the draft year 4 Success Criteria for report and recount writing].
2. **Carl:** I wonder if in end of year 2, ’cause it says in those Literacy [Learning] Progressions, that it’s can spell most of the words in essential lists level 1 and 2 and some of 3 and 4.
3. **Kerry:** We’ve got “Spells most high-frequency words” [reading from the draft year 2 Success Criteria for report and recount writing], would we rather change that?
4. **Carl:** I wonder if it [the year 2 criterion] should be essential lists 1 and 2.
5. **Kerry:** It’s a wee bit tighter isn’t it? . . . So instead of “spells most high-frequency words correctly” it will say “spells essential lists 1 and 2 correctly.”

(OBS: CHS, 5, SW, P. 17, L. 3–28)

During this negotiation, Carl used his school’s system of referring to the centrally developed writing standards (lines 4–7) to justify tightening the specification of the year 2 spelling criterion (lines 11–12). Thus, he drew attention to the lack of distinction between the existing year 2 (line 8) and year 4 (line 1–2) spelling criteria. This resulted
in the development of a clearer criterion for measuring students’ spelling at the year 2 level (lines 13–15). Through his involvement in the clarification of this criterion, Carl demonstrated that he had the ability to make use of the shared repertoire that he and his colleagues had developed to guide their moderation practice. For Carl, having access to this repertoire enabled him to contribute to the development of a more effective success criterion.

Through no fault of Carl’s, however, the revision he proposed was accidentally recorded as a year 1 rather than a year 2 criterion (see Appendix B, Rubric 2). Unfortunately, this error was not detected during subsequent revisions of this school’s sets of Success Criteria. As a result, this criterion appeared as a year 1 descriptor on the Success Criteria that were used to inform both the teaching and assessment of student writing during terms 2, 3, and 4 (see Appendix B, Rubrics 2–4). This error meant that the writing of a group of students was not being measured in relation to nationally agreed expectations. As such, this error posed a threat to the validity of the teachers’ interpretations of student writing at this level.

Overall, the data that have been presented in this section demonstrate that using involvement in moderation for the development of local assessment criteria can strengthen teachers’ assessment capability. Specifically, these data have shown how the teachers at Central Heights School were able to use their participation in moderation to get better at developing locally defined success criteria. In this regard, it has been noted that these teachers benefited from having the opportunity to work alongside a colleague who had a robust understanding of the properties of effective success criteria.

**Developing a shared perspective: Making connections between assessment and identifying next teaching and learning steps**

At Central Heights School, taking part in moderation also enabled teachers to get better at using their judgement-making processes for identifying next teaching and learning steps. In this way, these teachers used their participation in moderation to strengthen their ability to engage in what New Zealand’s Ministry of Education (2007) would describe as “learning inquiry” (p. 35). Forming part of the larger teaching as inquiry cycle, the learning inquiry process is expected to enable teachers to identify both how their teaching shaped student learning and what the implications are for future teaching. For these reasons, the learning inquiry process is promoted as a mechanism for strengthening both teaching and learning (Ministry of Education, 2007). The data that are presented within this section show how teachers at Central Heights School used
their involvement in moderation to develop a judgement-making procedure that brought the identification of next teaching and learning steps to the fore. It is argued that engaging in these processes enhanced the teachers’ ability to use assessment information for inquiring into and improving their teaching.

In order to use involvement in moderation for getting better at identifying next teaching and learning steps, the teachers at Central Heights School developed a system of using their Success Criteria to inform their judgements of student achievement. This approach became part of the shared repertoire that the teachers utilised to inform their moderation practice. Table 6 provides details of the number of direct references that the teachers at Central Heights School made to their locally developed Success Criteria during each of the observed focus-syndicate and full-staff judgement-making sessions. A direct reference to the Success Criteria involved a teacher either reading part or all of a criterion aloud or prompting a colleague or colleagues to make use of this resource.

Table 6

<table>
<thead>
<tr>
<th>Meeting number</th>
<th>Meeting type</th>
<th>Number of direct references to Success Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Focus syndicate</td>
<td>63</td>
</tr>
<tr>
<td>4</td>
<td>Full staff</td>
<td>63</td>
</tr>
<tr>
<td>6</td>
<td>Focus syndicate</td>
<td>53</td>
</tr>
<tr>
<td>7</td>
<td>Full staff</td>
<td>38</td>
</tr>
<tr>
<td>9</td>
<td>Focus syndicate: paired</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>Focus syndicate: paired</td>
<td>54</td>
</tr>
<tr>
<td>11</td>
<td>Full staff</td>
<td>24</td>
</tr>
</tbody>
</table>

At Central Heights School, the practice of using their agreed assessment criteria to inform discussions about student writing helped the teachers to ensure that their conversations remained focused on the work that they were appraising. Engaging in this process also enabled the teachers to share and reflect upon the inferences that had informed their judgements of student writing. Given that the validity of an assessment can be thought of as a measure of the quality of the inferences and decisions that are
made on its basis (Messick, 1989), participating in this process helped the teachers to validate their judgements of student writing. This approach was established during the senior syndicate’s first judgement-making session (see Appendix E, Meeting 3). The exchange below, which took place during this meeting, occurred after Kerry questioned Chris’s judgement that a piece of student writing had fulfilled their year 5 Success Criteria:

1  **Chris**: She’s using complex sentences [referring to the first year 5 success criterion for personal experience writing].
2  She’s punctuating, it’s pretty accurate [referring to the second year 5 criterion]. I mean the only thing was speech, she’s using exclamation marks, speech marks. She forgot a comma in one of them. . . .
3  **Kerry**: Involves the reader in the experience [referring to the final year 5 criterion]? She sort of does.
4  **Margaret**: I think that she does. I mean I was right there with her. . . .
5  **Chris**: Adds detail to it [referring to the sixth year 5 criterion], yeah. She describes the dog. . . .
6  **Kerry**: She does well on that. . . .
7  **Chris**: I thought for the next one [the year 6 level], she didn’t [achieve it]. ’Cause I sort of looked . . . but . . . there’s no editing and she’s not paragraphing. There’s a few things there that she wasn’t doing, so I thought it was comfortable there [at the year 5 level].

(BOSS: CHS, 3, FS, P. 39, L. 29 – P. 40, L. 7)

Within this conversation, the teachers used their Success Criteria to systematically review this student’s writing. As well as providing Chris with the opportunity to justify his judgement-making process (lines 1–6, 11–12, & 14–18), engaging in this discussion also allowed Kerry (lines 8 & 13) and Margaret (lines 9–10) to contribute to the validation process.

This system, of using their Success Criteria to inform their judgements of student work, also enabled the teachers at Central Heights School to use these conversations to identify next teaching and learning steps. This is a practice that a number of the longer-serving teachers at Central Heights School might have been
introduced to during their school’s participation in the Ministry of Education’s AToL project (Poskitt & Taylor, 2008). As previously explained, Central Heights School’s involvement in this initiative occurred at least 7 years prior to this study’s data-collection phase. Taking part in moderation provided all teachers at this school with the opportunity to refocus their attention on using judgement-making conversations for identifying next teaching and learning steps. As was the case with the development of other moderation systems at this school, Kerry played an important role in establishing this practice. Indeed, it was an approach that she began promoting as soon as she and her colleagues had finished drafting their first set of Success Criteria. Specifically, as their initial full-staff moderation meeting (see Appendix E, Meeting 2) was about to conclude, Kerry announced,

The other thing that I was going to suggest is that these [Success Criteria] are only going to be useful if they’re fed back to kids, if they’re useful bits of feedback to kids. . . . When you’re giving feedback, just choose one of those [Success Criteria] that you are pushing with that child to feed back in the next learning step sort of thing. And that should be helpful to us when we are making our feedback specific to the child’s writing.

(OBS: CHS, 2, SW, P. 26, L. 18–25)

Initially, Kerry needed to prompt some of her colleagues to use their involvement in judgement making for the identification of next teaching and learning steps. For example, during this school’s first full-staff judgement-making session (see Appendix E, Meeting 4) Kerry encouraged Rose, a junior-syndicate colleague, to use the evaluation of a student’s writing for this purpose. Notably, Rose was the first teacher to share student writing during this judgement-making session. In this case, the student whose work was being reviewed was in year 1 and had been at school for just 9 weeks:

1  **Rose:** We could talk about the negatives. The positives are that she’s circling words that she is not sure about. She is, she writes “saw” all the time, so now she knows to write it correctly. And she’s got other basic words that she knows as well.
2  **Kerry:** So, Rose . . . instead of us talking about negatives, what would her next learning step be for you?
Rose: Well, I’m carefully pushing for her, and it will be [child’s name] next step, the finger spaces, the circling as she goes, because she wants to get it done very quickly. We’ll work on probably neatening up. But more importantly, the deeper features will be the words that she is [unsure of]. While she’s done a very good job of “supermarket,” sometimes she gets a bit pooped. And so we want some more expressive words to go in her stories.

(OBS: CHS, 4, SW, P. 2, L. 48 – P. 3, L. 9)

Here, Kerry’s prompt had a distinct effect on the way in which Rose appraised her student’s work. Prior to Kerry’s question, Rose’s evaluation appeared to be focused on identifying the success criterion that this student’s writing had fulfilled. Specifically, her opening comments (lines 1–5) about the student’s ability to spell “saw” and other basic words correctly indicate that Rose considered this child was beginning to meet the after-6-month surface-feature criterion “can record some high-frequency words” (as shown in Appendix B, Rubric 1). After Kerry’s question, Rose’s approach shifted and she began to use her involvement in judgement making to identify this student’s next learning steps. Initially, she focused on identifying those skills (lines 9–11), such as leaving spaces between words and writing more neatly, that this child would need to acquire to fulfil the surface-feature criterion “can read back what they have written” (as shown in Appendix B, Rubric 1). Later, her focus shifted to this student’s command of deeper feature skills. Her comments here indicate that she had linked attainment of the deep-feature criterion “can record personal experiences” (as shown in Appendix B, Rubric 1) with the ability to attempt unknown words (lines 12–15). Notably, the latter skill is associated with the surface-feature criterion “can record the initial or dominant sounds in words” (as shown in Appendix B, Rubric 1). With Kerry’s assistance, Rose was able to reorient her judgement-making approach to focus on the identification of this student’s next learning steps. Such an approach enables teachers to use judgement making as an opportunity to inquire into the effects of their teaching and consider the

---

8 Because this exchange occurred at the beginning of this school’s initial full-staff judgement-making session, the teachers had not formalised their judgement-making procedure. Therefore, although the following analysis shows how Rose’s appraisal was linked with her school’s Success Criteria, it should be noted that such references to the Success Criteria are not sufficiently explicit to have been included in Table 6.
implications for future instruction. As such, it is consistent with the New Zealand Ministry of Education’s (2007) conception of effective pedagogical practice.

By the end of the year, the practice of using moderation sessions for identifying next teaching and learning steps had become a reflexive and valued feature of these teachers’ judgement-making processes. This was evident, for example, when two members of the senior syndicate met during the third term for a paired judgement-making session (see Appendix E, Meeting 9). Within this meeting, William and Margaret discussed their judgements of student responses to a common assessment. For this task, students had been asked to write a character description of an athlete who was competing at the ancient Olympic games. Although Kerry was not present at this paired moderation session, these teachers identified next learning steps for five of the six students whose writing they appraised. The following exchange, which occurred as Margaret and William discussed their judgements of a year 4 student’s writing, demonstrates how these teachers had integrated the identification of next learning steps into their shared repertoire:

1 Margaret: “Can describe the character’s behaviours specifically” [reading from the year 4 Success Criteria]. So I think he’s done that. I would have liked him to have done a bit more.
2 William: He’s started off really well. Like you said about his [character’s] first [discus] throw, he got [placed] third. Like he [the student] really set himself up well . . .
3 Margaret: He’s been specific but he just hasn’t been, he hasn’t elaborated enough . . .
4 William: That could be his next step . . . just extending that [description of the character’s behaviour] so he goes further.
5 Margaret: A bit more, yeah.
6 William: A bit more detail and just being a bit more specific.

(OBS: CHS, 9, FS, P. 7, L. 45 – P. 9, L. 2)

Here, the use of two complementary judgement-making systems enabled these teachers to use this conversation to identify a next learning step. As previously explained, the teachers at this school had developed a system of using their Success Criteria to inform their judgements of student work. This was linked with a second, related system, which
involved looking for evidence of the extent to which each student’s writing had fulfilled
the Success Criteria. Within this exchange, Margaret employed the former approach as
she initiated the discussion about this student’s work (lines 1–2). William then utilised
the latter approach as he drew on an example from the student’s writing (lines 5–7).
This allowed him to engage with and respond to Margaret’s initial appraisal (lines 2–4).
Crucially, it was the use of these systems that allowed the teachers to both identify and
reach agreement about a next learning step for the student concerned (lines 10–14).

As the following exchange reveals, these teachers recognised the value of using
their judgement-making sessions for systematically identifying next learning steps for
students. This was a perspective that they both expressed during the post-moderation
interview that they participated in. This took place the day after their paired judgement-
making session:

1    Margaret: And even after just doing, how many [samples]
2            did we do [discuss] yesterday, 3 or 4 each? . . . You could
3                  see where there were things, oh well I know what I’ll be
4                doing next. . . .
5    William: Right, that’s the [student’s] next learning step.
6            So, I obviously need to change it [my teaching] and think,
7                  right, with my language programme.
8    Margaret: So that becomes a teaching point doesn’t it. . . .
9    William: That’s where I’ve got to do my next teaching
10            step as well, not just the kids’ learning step. Well they need
11            that teaching first, don’t they?

(INT: CHS, 1, FST, GRP, P. 5, L. 10–46)

During this exchange, William and Margaret reflected on their paired judgement-
making session (lines 1–4) and linked the process of identifying next learning steps
(line 5) with inquiring into and making adjustments to their teaching (lines 2–4 & 5–11).
Their comments suggest that involvement in moderation enhanced their understanding
of the ways in which student assessment information can be used formatively for
inquiring into and improving teaching and learning programmes. Notably, it was this
aspect of their practice that appeared to excite these teachers the most about moderation.
As explained earlier, effective professional learning initiatives typically provide
participants with a compelling reason for engaging in the requisite learning (Timperley
et al., 2007). These teachers sustained their involvement in moderation because they understood it could help them identify next teaching and learning steps.

There were strong indications that other teachers at this school shared William and Margaret’s perspective about the benefits of using their involvement in moderation practice for this purpose. For example, when asked during an interview to identify the main purposes of the moderation activities he was involved in, Chris replied,

1. Knowing where the kids are at. Just to know where they are
2. at and what the next step is. ’Cause I mean, I suppose it’s
3. like any progression . . . going well they [the students] know
4. that, what can they do next? How do they move on? I guess
5. if you [the teacher] don’t know where they are, where do
6. you start from?

(INT: CHS, 2, FST, IND, P. 2, L. 35–39)

Like his senior-syndicate colleagues, Chris linked the process of identifying his students’ next learning steps (lines 1–2) with obtaining the information he required to inform his teaching decisions (lines 4–6). The identification of next teaching and/or learning steps also featured prominently in the responses that teachers at this school provided to an open-ended questionnaire item about the main purposes of the moderation activities in which they were involved (see Appendix K, Q. 21). Table 7 provides details of the responses that teachers across the three schools provided to this question. Although the teachers at Central Heights School reported that moderation served a variety of main purposes, the identification of next teaching and/or learning steps was the only function that over half of the respondents at this school mentioned in their answers. This provides an indication of the value that these teachers attributed to this aspect of their moderation practice.
Table 7

*The main purposes of moderation identified at each participating school*

<table>
<thead>
<tr>
<th>Identified purposes</th>
<th>Central Heights School</th>
<th>Greenville School</th>
<th>Riverside School</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>(n = 9)</em></td>
<td><em>(n = 6)</em></td>
<td><em>(n = 4)</em></td>
</tr>
<tr>
<td>Improving or ensuring consistency</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Reporting and assessing</td>
<td>1</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>Increasing teachers’ assessment confidence</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Strengthening teachers’ assessment capability</td>
<td>2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Identifying next teaching and/or learning steps</td>
<td>5</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Improving student outcomes</td>
<td>—</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>Appreciating children’s writing</td>
<td>1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Developing an understanding of the learning progression in writing</td>
<td>4</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

The data presented in Table 7 also reveal that Central Heights School was the only school at which the identification of next teaching and/or learning steps was perceived as a main purpose of moderation. This was a view of moderation that the principal of Central Heights School also advocated. In fact, when asked to identify the main purpose of moderation she replied, “I see it [the purpose of moderation] as like teacher inquiry” (INT: CHS, 4, PL, IND, P. 2, L. 10–11). She later elaborated on this perspective and explained,

1. We are [through moderation] honing the teachers’ thinking
2. really, around … how that child presents. . . . And the
3. knowledge of the child is just huge too in a teaching and
4. learning situation. If you are going to lift achievement, the
5. knowledge of the child and the accurate assessment to make
6. a [learning] goal to shift [that child] is really, really
7. important. So I think the whole school, each child, will
Like her colleagues, the principal of Central Heights School linked involvement in moderation with the identification of next teaching and learning goals (lines 4–7). This, coupled with the sense that the identification of these goals would enable teachers to improve student outcomes (lines 4–7), provided her with a sufficiently compelling reason for valuing and supporting her school’s moderation processes (lines 7–9). The development of this school-wide shared perspective about the purpose of moderation provides further evidence of what Wenger (1998) would describe as the negotiation of a joint enterprise. For the principal and the teachers at Central Heights School, this joint enterprise provided them with the rationale they required to sustain their mutual engagement in moderation.

At Central Heights School, participating in the development and use of a shared repertoire – that included local assessment criteria and common judgement-making systems – enabled the teachers to arrive at a shared perspective about the purpose of their involvement in moderation. The teachers at this school came to understand that participating in these processes allowed them to inquire into and strengthen their teaching. This is important because the New Zealand Ministry of Education (2011c) has linked teachers’ involvement in such inquiry processes with the promotion of student learning. At Central Heights School, the shared perspective that teachers held about the purpose of moderation was closely aligned with the vision that both the principal and the deputy principal had for their school’s involvement in these processes. Because this school’s deputy principal, Kerry, encouraged her colleagues to use their judgement-making conversations for identifying next teaching and learning steps, these teachers were able to use their involvement in moderation to strengthen this facet of their assessment capability.

**Chapter conclusion**

This chapter has identified the resources, systems, and perspectives that enabled teachers to learn about assessment for learning principles and practices through their involvement in moderation. Building on Wenger’s (1998) conception of practice, it has highlighted the ways in which the development and use of tangible resources, common systems, and shared perspectives shaped the learning opportunities that involvement in
moderation afforded teachers. Likewise, it has identified the ways in which schools utilised their existing assessment expertise to maximise the professional learning benefits that teachers garnered from participating in moderation. The data presented within this chapter have shown how teachers were able to use their involvement in the conceptualisation, development, and implementation of moderation processes to strengthen their assessment for learning capability. The next chapter builds on these findings and explores the opportunities that involvement in moderation provided teachers for learning about factors that can affect the dependability of assessment information.
Chapter 5. Moderation and the negotiation of meaning: 
Conversations about standards, judgement making, and assessment tasks

Teachers require sound assessment information to enable them to inquire into and adjust their teaching programmes (Ministry of Education, 2007, 2011c; Newton, 2010). This chapter reports on the ways in which teachers used their involvement in moderation to strengthen their ability to engage in this important inquiry process. It examines how and what teachers’ experiences of moderation taught them about factors that can affect the dependability of assessment information. Here, Wenger’s (1998) conception of meaning, or “learning as experience” (p. 5), is used to explore how involvement in moderation shaped the assessment capability of the participating teachers. The opening section examines the ways in which taking part in moderation informed teachers’ understandings of standards and assessment criteria. Next, the focus shifts to judgement making. Here, attention is paid to the ways in which involvement in moderation shaped teachers’ awareness of those factors that can affect the dependability of qualitative judgements of student achievement. The chapter concludes by exploring what teachers’ experiences of moderation taught them about designing, selecting, and administering assessment tasks. As well as examining what the participating teachers learned about assessment from their experiences of moderation, this chapter also investigates how the leadership of each school’s moderation processes affected teacher learning. These findings indicate that involvement in school-based moderation processes can enable teachers to strengthen their assessment capability in ways that should also result in the production of more dependable student assessment information. Yet these findings also suggest that to have this effect, moderation processes need to be carefully planned and coherently led.

Understandings of standards and assessment criteria

Taking part in moderation processes provided the teachers at both Central Heights and Greenville schools with opportunities to clarify their understandings of the centrally developed writing standards (Ministry of Education, 2009c, 2010b). These experiences also enabled them to establish common understandings of their local assessment criteria. Such shared understandings are important because they inform not only the decisions that teachers make about selecting and devising assessment tasks (Darr, 2005b) but also
the judgements of student achievement at which they arrive (Klenowski & Wyatt-Smith, 2010a, 2014; Maxwell, 2010). For these reasons, the development of shared understandings of standards and assessment criteria was identified in the literature review as a factor that can affect the dependability of student assessment information. The following section begins by describing those experiences of moderation that enabled teachers to develop shared, syndicate-level understandings of assessment criteria. It then examines how and why the teachers at Central Heights School were able to use their involvement in moderation to develop school-wide understandings of their Success Criteria.

**Syndicate-level understandings**

For the teachers at both Central Heights and Greenville schools, the opportunity to co-construct assessment criteria during syndicate meetings resulted in the development of shared, syndicate-level understandings of these criteria. Because these teachers used the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to inform their local assessment criteria, taking part in these processes also enabled them to clarify their understandings of the standards that are specified within these centrally developed documents. The findings that are presented next demonstrate how involvement in the practice of criteria development enabled these teachers to engage in a process that Wenger (1998) has described as “the negotiation of meaning” (p. 52). According to Wenger, new meanings are negotiated when members of a community of practice engage in acts of participation and reification. Although Wenger’s application of the term participation is consistent with common use, his treatment of reification requires further explanation. He described this concept as “the process of giving form to our experience by producing objects that congeal this experience into ‘thingness’” (p. 58). For Wenger, the notion of reification can encapsulate both a process and its product. Within the context of moderation, for example, the concept of reification might be used to refer both to a process that was developed for producing school-specific assessment criteria and to the concrete products that this process rendered.

At both Central Heights and Greenville schools, the practice of using syndicate meetings to participate in the co-construction, or reification, of assessment criteria enabled the teachers in each focus syndicate to negotiate shared understandings of their local achievement expectations. At Greenville School, a number of instances of this occurred as the teachers in the middle syndicate developed Must, Should, Could charts for students in years 2 and 4 (see Appendix C). As explained in Chapter 4, these charts
were written in child-friendly language and specified what student writing must, should, and could exhibit by the end of a given year of schooling. During the meeting within which the teachers developed these charts (see Appendix F, Meeting 3) they made 74 references to the Literacy Learning Progressions (Ministry of Education, 2010b) and negotiated how to translate these criteria into “kid speak” (OBS: GS, 3, FS, P. 13, L. 14).

The following exchange provides an example of how this participative approach helped the teachers to align their understandings of achievement expectations. Within this dialogue, the year levels of the students in each teacher’s class are indicated after their name:

1 Phillip [Y4]: [Reading from Literacy Learning Progressions: After 2 years at school] Compose “mainly simple and compound sentences, with some variation in their beginnings.” . . . I would put, you could . . . use different sentence beginnings.
2 Sophia [Y2 & Y3]: I would say should.
3 Susan [Y3 & Y4]: Should, yeah. I think so. Yeah.
4 Phillip [Y4]: Oh, you reckon should?
5 Sophia [Y2 & Y3]: I would expect mine to be using them as a should. They should not be doing “and then, and then and then.”
6 Phillip [Y4]: Year 2 is the year level that I have got the least experience with.


This dialogue captures the way in which the teachers utilised both the centrally developed writing standards (Ministry of Education, 2010b; lines 2–4) and their practitioner knowledge (lines 4–5, 6–7, & 9–11) as they collaboratively co-constructed their school-specific assessment criteria. In this conversation, the teachers engaged in a negotiation about whether year 2 students should be expected to vary their sentence beginnings. This skill, which the centrally developed Literacy Learning Progressions (Ministry of Education, 2010b) state students are expected to demonstrate in “some” (p. 13) of their sentence beginnings at the year 2 level, becomes an unqualified expectation after 3 years at school. Because Phillip had limited experience working with year 2 students (lines 12–13), his expectations were lower than those of his colleagues. Notably, both Susan and Sophia, who had greater familiarity with teaching at this level,
articulated expectations that were more closely aligned with those specified in the centrally developed Literacy Learning Progressions. For Phillip, the opportunity to co-construct assessment criteria with colleagues who had greater experience working with year 2 students served not only to validate the centrally developed expectations but also to ensure that a shared local understanding of this expectation was established.

Like their counterparts at Greenville School, the teachers at Central Heights School utilised both the centrally developed writing standards (Ministry of Education, 2009c, 2010b) and their practitioner knowledge to negotiate shared understandings of local assessment criteria. This was an approach that the teachers at both of these schools used as they engaged in negotiations about the time by which students should demonstrate mastery of those skills that the centrally developed writing standards introduce progressively. Like the use of variation in sentence beginnings, the use of complex sentences is gradually introduced. This is illustrated in Table 8, which reproduces the expectations about the use of complex sentences that are described in these documents.

Table 8

<table>
<thead>
<tr>
<th>Year Level</th>
<th>Literacy Learning Progressions</th>
<th>Reading and Writing Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>After 3 years at school</td>
<td>Attempting to write complex sentences</td>
<td>Attempts at some complex sentences</td>
</tr>
<tr>
<td>By the end of year 4</td>
<td>Using mainly simple and compound sentences, along with some complex sentences, that vary in their beginnings, structures, and lengths and are mostly correct grammatically</td>
<td>Attempts at complex sentences</td>
</tr>
<tr>
<td>By the end of year 6</td>
<td>Using some complex sentences that are mostly correct grammatically</td>
<td>Some complex sentences that are mostly correct grammatically</td>
</tr>
</tbody>
</table>
At both Central Heights and Greenville schools, the teachers utilised their practitioner knowledge to determine when they should expect students to include complex sentences in their writing. Several examples of this occurred as the teachers in Central Heights School’s senior syndicate negotiated which sentence structures year 4 students should be expected to write accurately (see Appendix E, Meeting 1). In keeping with this school’s practice of using the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to inform their local Success Criteria, these teachers first reviewed the year 4 expectations that were specified within these documents. Following this, Chris, a teacher who had 8 years’ experience working with students in years 5 and 6, commented, “So, I doubt we’ll get complex sentences [at year 4]” (OBS: CHS, 1, FS, P. 4, L. 27–8). Kerry, who had taught students in years 5 and 6 for 15 years, later echoed this sentiment. She reflected,

I don’t think that many of our year 4 kids are doing that, not from what I have seen. . . . I think we will leave that, the complex sentences. That is what we are aiming for at year 6 I think.

(OBS: CHS, 1, FS, P. 7, L. 15–41)

In response to this Chris noted, “Well, it [complex sentences] still is in the year 6 one [referring to the Literacy Learning Progressions]” (OBS: CHS, 1, FS, P. 7, L. 45–6). These teachers subsequently classified the inclusion of complex sentences as a “beginning to” skill at the year 4 level (OBS: CHS, 1, FS, P. 8, L. 25) and agreed that it should first appear as an expectation in their year 5 Success Criteria (see Appendix B, Rubric 1). This negotiation enabled the teachers within this syndicate to reach a common understanding about how to evaluate this aspect of student writing.

A similar negotiation occurred at Greenville School when the teachers in this school’s middle syndicate discussed when students should be expected to include complex sentences in their writing (see Appendix F, Meeting 3). The two teachers who actively participated in this discussion were, at the time of data collection, teaching students at the year 4 level. Between them, they had a combined total of 19 years’ experience teaching at this level. Although these teachers felt that students must be including compound sentences in their writing at year 4 (see Appendix C), they agreed that it was more appropriate to categorise the inclusion of complex sentences as a skill that students should be demonstrating at this year level (see Appendix C). The teachers in Greenville School’s middle syndicate viewed those skills that they categorised at each year level as a must as the things that students had to demonstrate to meet the
standard for that year level. In contrast with this, those skills that were included at each year level in the *could* category were intended to provide students with the opportunity to work at the year level above. Therefore, like their colleagues at Central Heights School, these teachers perceived that the inclusion of complex sentences was a skill that students were only beginning to demonstrate at the year 4 level.

For the teachers at both Central Heights and Greenville schools the development, or reification, of local assessment criteria was an interpretative process. The findings that have just been presented show how these teachers drew on their practitioner knowledge to clarify their understandings of the centrally developed writing standards (Ministry of Education, 2009c, 2010b) and co-construct local assessment criteria. Crucially, it was through engaging in these shared interpretive acts that the teachers developed common syndicate-level understandings of their local assessment criteria. These common understandings, which reflected each syndicate’s history of mutual engagement in moderation, became an important part of their shared repertoire. Wenger (1998) has emphasised that a community’s shared repertoire is useful because it becomes “a resource to be used in the production of new meanings” (p. 83). This idea is explored in the following section, which shows how the teachers at one school used their involvement in moderation to establish common school-wide understandings of their local assessment criteria.

**School-wide understandings**

At Central Heights School, involvement in full-staff moderation meetings provided teachers with opportunities to negotiate shared, school-wide understandings of their local Success Criteria. These opportunities occurred when teachers met as a full staff to review their local criteria. They also occurred during full-staff judgement-making sessions. The following example shows that taking part in criteria-development sessions that involved the full staff provided teachers with opportunities to negotiate common school-wide understandings of their local assessment criteria. Meetings of this type typically occurred after the teachers had met in their syndicates to draft Success Criteria for their own teaching levels. At this school, the skills required for all writing purposes were detailed in the surface features section of their Success Criteria. Conversely, those associated with a specific genre or purpose were specified in the deep features section.

The negotiation that the upcoming example centres around occurred during a full-staff meeting at which the teachers reviewed their Success Criteria for personal experience writing (see Appendix E, Meeting 2). This negotiation lasted almost 4½
minutes and involved over half of those present (n = 15). Participants in this discussion included the principal, a special needs teacher, and teachers from both the junior (JNR) and senior (SNR) syndicates. This conversation began after a junior-syndicate teacher shared the Success Criteria that she and her colleagues had previously drafted. During the ensuing conversation, participants sought to clarify whether the term *clearly* in the draft criterion “expresses ideas clearly” related to syntax or semantics (OBS: CHS, 2, SW, P. 6, L. 22). The teachers in the junior syndicate had classified this as a deep-feature criterion to indicate that it described an expectation that they associated specifically with personal experience writing. This discussion began when Kerry asked, “What would you mean by *clearly*?” (OBS: CHS, 2, SW, P. 6, L. 31). Her query initiated the following exchange:

1 **Rose (JNR):** Because again some children might say “I shop” or “I like that” and we want them to be able to say “I like going to the shop. I like eating my ice-cream.” And hopefully that comes across more clearly.

2 **Kerry (SNR):** Isn’t that to do with writing an accurate sentence?

3 **Judy (JNR):** It depends, with the “accurate” though, whether you are thinking about the linguistic side or whether you are thinking about the writing that grammatical side to the linguistic side, or the process of actually writing. (OBS: CHS, 2, SW, P. 6, L. 33–46).

Within this discussion, Kerry challenged her junior-syndicate colleagues to explain how the criterion under review might measure more than just a sentence’s structural accuracy (lines 5–6). This reflected the understanding that she had gleaned from Rose’s explanation (lines 1–4). Judy, however, had a different view. Drawing on the shared understanding that she and her junior-syndicate colleagues had forged when they developed this criterion, she attempted to acknowledge the distinction that can exist between semantic and syntactic accuracy (lines 7–10). Yet the teachers continued to talk around this topic for almost another minute before a comment from the principal enabled Judy to make her point more effectively:
Here, it was the principal’s introduction of the term meaningful (line 2) that helped Judy to articulate that she had linked the original criterion with semantics rather than syntax (line 5). Following Judy’s clarification, the teachers deliberated over this issue for some time. Indeed, during the ensuing 2½ minutes their conversation moved back and forth between the roles of syntax and semantics until Kerry proffered a solution:

Within this exchange, Kerry helped her colleagues to develop shared understandings of two Success Criteria. The first instance of this occurred when she demonstrated that an existing criterion (lines 4–6) addressed the need that the teachers had identified in relation to measuring sentence structure, or syntax (lines 2–3). In this way, she linked their earlier conversations about syntax with this criterion. In doing so, she imbued this description of achievement with new, commonly understood meaning. Likewise, Kerry played a critical role in enabling her colleagues to reach a common understanding of the deep-feature criterion that had initially prompted their lengthy negotiation. Specifically,
it was Kerry’s revision of the deep-feature criterion (lines 6–8) that her colleagues finally agreed (lines 9–13) reified the sentiment that they wanted to convey.

Because the teachers at Central Heights School used their Success Criteria to inform their judgements of student writing, taking part in judgement-making sessions that involved the full staff provided them with additional opportunities to negotiate shared, school-wide understandings of these descriptions of achievement. The following example shows how participating in such a judgement-making session (see Appendix E, Meeting 7) enabled Chris to deepen his understanding of the criteria that described the writing behaviours that students were expected to display after 6 months at school. Having listened to a colleague read the writing of a student who had been at school for about 13 weeks, Chris, who taught students in years 5 and 6, asked,

Can they [the students] do that unaided? Like do they do it [write like that] after a term and a bit [at school]? . . . Are they actually given words? . . . I don’t know how it works in the juniors.

(OBS: CHS, 7, SW, P. 3, L. 15–24)

During the ensuing discussion, Chris’s junior-syndicate colleagues drew on evidence from their students’ writing and used descriptions of the classroom environment to explain how their students were able to fulfil a number of the “after 6 months” criteria. For example, Rose explained how working alongside a child who had been able to record both the “l” and the “k” in the word like had provided her with evidence that he was beginning to achieve the criterion “can record the initial or dominant sounds in words” (as shown in Appendix B: Rubric 2). Similarly, Kath explained how “having words . . . around the walls and on their desks and things” (OBS: CHS, 7, SW, P. 3, L. 21) helped the students to attain the criterion “can record some high-frequency words” (as shown in Appendix B: Rubric 2). Being provided with these tangible examples seemed to help Chris negotiate a richer understanding of each criterion.

At the conclusion of another full-staff judgement-making session (see Appendix E, Meeting 11), Jean and William talked about how taking part in this meeting had broadened their understandings of their school’s Success Criteria. Within this meeting, Kerry had asked the teachers to work with a colleague who taught in a different syndicate. Accordingly, William, who taught in the senior syndicate, paired up with Jean, a junior-syndicate teacher. To conclude this judgement-making session, Kerry asked her colleagues for some feedback about the experience of working in this way.
Jean and William’s responses provide an indication of the meanings that this experience enabled them to negotiate:

1. **Jean:** Well, it focuses you in on another [year] level that you wouldn’t really normally work with. And ’cause our jump is a couple of years over, it’s actually quite good. Because you kind of know the next one [year] up but not necessarily the ones [years] above that again. Or maybe, for you guys [in the senior syndicate], [the years] below.

2. **William:** It makes you look at the whole grid [set of criteria] and become more familiar. So instead of me just being . . . there [looking at the year 4 criteria], I’m . . . looking at all of it [and] sort of becoming more familiar.


Within their responses, both Jean (lines 1–2 & 3–5) and William (lines 7–10) talked about how working with a colleague from another syndicate had enabled them to extend their understandings of the Success Criteria beyond those specified at their own instructional level.

The teachers at Central Heights School were able to negotiate and sustain shared, school-wide understandings of their Success Criteria because they had regular opportunities to participate in both syndicate-level and full-staff moderation sessions. The findings reported in Figure 1 show that the teachers at this school were the only participants in this study to have both regular and ongoing opportunities to participate in each of these types of moderation meeting.
Figure 1. The number of syndicate and full-staff moderation meetings that teachers participated in each term, by school.⁹ CHS = Central Heights School; GS = Greenville School; RS = Riverside School.

Wenger (1998) explained that the “mutual engagement of participants” (p. 73) is one of the dimensions that provides a community of practice with its coherence. In this regard, he emphasised that the factors that enable this engagement are a crucial feature of any practice. At Central Heights School, two factors contributed to teachers having regular opportunities to mutually engage with each other through their involvement in moderation. These were the existence of a clearly defined moderation plan and a commonly understood moderation leadership system.

This school’s moderation plan (see Appendix O), which was linked with their professional development goals for literacy, had been developed by the full staff before teachers began their moderation work. Although it was modified slightly during the year, this plan provided the blueprint for Central Heights School’s moderation activity. Because the moderation plan specified that moderation meetings would occur at both the syndicate and the full-staff level, opportunities for teachers to expand their understandings of local assessment criteria beyond their teaching level were built into this school’s moderation approach. Additionally, the plan also clearly stated that moderation activities were expected to be both frequent and ongoing. Specifically, it

⁹ Within this figure inter-syndicate meetings are categorised as syndicate meetings.
stated that teachers would be involved in at least three moderation meetings per term, two of which would be devoted to criteria development. It also specified that this pattern of activity would repeat each term. This is important because evidence-based research (Timperley et al., 2007) has identified that those professional learning initiatives that are associated with positive outcomes typically provide teachers with frequent learning opportunities over an extended timeframe.

Although not explicitly stated in this plan, it was also understood that Kerry, who was this school’s deputy principal, had overarching responsibility for leading these activities. In this capacity, Kerry was observed facilitating all the full-staff and senior-syndicate moderation sessions that occurred at Central Heights School (see Appendix E). In fact, Kerry’s leadership was something that a number of teachers chose to comment on when responding to a questionnaire item about the role their principal played in their school’s moderation processes. For example, Jean explained, “Our DP [deputy principal] leads all moderation staff meetings” (QST: CHS, 3, Q. 22). Similarly, Rose commented, “[the principal] leaves [the] DP [deputy principal] to use her expertise to facilitate [these processes]” (QST: CHS, 8, Q. 22). In her role as the leader of this school’s moderation activities, Kerry had the full support of the principal, who attended and actively participated in all but two of her school’s full-staff moderation meetings (see Appendix E).

A somewhat different situation existed at Greenville School. Here the teachers in the middle and senior syndicates used their initial moderation meetings (see Appendix F, Meetings 1 & 2) to begin planning and conceptualising their school’s moderation approach. Susan, the leader of this school’s middle syndicate, spearheaded these efforts. During these meetings, the teachers in the middle and senior syndicates worked together and agreed to create sets of school-specific assessment criteria. The teachers subsequently developed these criteria, which they referred to as Must, Should, Could charts, during separate syndicate-level (as opposed to inter-syndicate) meetings (see Appendix F, Meeting 3). Details about the shared understandings that involvement

---

10 The referencing system that I use to cite open-ended questionnaire (QST) responses is as follows. There are three parts to the reference. First the school is identified as CHS (Central Heights School), RS: (Riverside School), or GS (Greenville School). Then the specific questionnaire that the response is drawn from is cited. In this case the “3” denotes that it was the third questionnaire that was collected at this school. Finally, the question that the response was associated with is specified. So, in this case the comment cited was a response to question 22.
in this process enabled the teachers in the middle syndicate to arrive at were provided earlier in this chapter.

After these charts had been created, Susan, who had facilitated all middle- and inter-syndicate moderation sessions, used a full-staff meeting to introduce this system to the junior-syndicate teachers (see Appendix F, Meeting 5). In this respect, Susan engaged in an activity that Wenger (1998) would describe as “brokering” (p. 109). He explained that the role of broker is typically assumed by the person(s) within an organisation with responsibility for a special project that spans functional units. Because Susan had assumed responsibility for spearheading Greenville School’s moderation activities (the special project), she needed to establish new connections across each of this school’s syndicates (the functional units). Wenger explained that such connections are required to enable co-ordination and to create opportunities for new meanings to be negotiated. The job of a broker is, therefore, a complex one. Without sufficient legitimacy, a broker will struggle to address conflicting interests and may not be able to influence the development of a new practice. In her capacity as a broker, Susan used this full-staff meeting to explain the Must, Should, Could system to her junior-syndicate colleagues. At the end of this session, the deputy principal, who was also the leader of the junior syndicate, agreed that she and her colleagues would develop a Must, Should, Could chart that described writing expectations for students in year 1.

Although the teachers in the junior syndicate did later meet to co-construct assessment criteria, they ultimately decided against using the Must, Should, Could system. This decision left some teachers feeling unsure about how to proceed. For example, Susan later commented,

The juniors . . . haven’t quite come on board with the Must, Should, Could [system] . . . We’ve got Donna, who’s a year 2 teacher . . . she just feels as if she’s drifting. Does she go with our . . . year 2 sheets? Does she go with what the juniors are doing? So I think we still need a bit of alignment there.

(INT: GS, 1, LT&FST, GRP, P. 43, L. 17–28)

It is also possible that the junior syndicate’s decision may have undermined Susan’s burgeoning sense that she was leading her school’s moderation approach. Following their decision to adopt an alternative moderation system, Susan did not attempt to schedule any further full-staff moderation sessions. When I later asked her whether there had been a formal understanding that she was leading her school’s moderation processes, Susan replied,
Well, not really, no. But I think I probably need to become a bit firmer about that. Because Christine [the deputy principal] said the other day, “oh well you’ve been leading this [moderation].” And I didn’t know that. . . . It’s just sort of happened.

(INT: GS, 1, LT&FST, GRP, P. 49, L. 18–25).

This statement provides some insight into the uncertainty that surrounded the leadership of Greenville School’s moderation processes. Although Susan had spearheaded her school’s moderation efforts and encouraged her colleagues to develop a school-wide moderation system, she did not see herself as the official leader of this initiative (lines 3–5). Despite the deputy principal’s assertion that she viewed Susan as the moderation leader (lines 2–4), her actions suggested otherwise. Specifically, her syndicate’s decision not to adopt the Must, Should, Could system appeared to undermine Susan’s attempts to broker a school-wide moderation approach. Regardless of its source, uncertainty existed at Greenville School about who had overarching responsibility for leading moderation. This may have played a role in limiting the opportunities that the teachers at this school were afforded for developing school-wide understandings of their assessment criteria.

For teachers at both Central Heights and Greenville schools, opportunities to co-construct assessment criteria with members of their teaching teams resulted in the development of shared understandings of these criteria at the syndicate level. Yet it was only at Central Heights School, where teachers also had opportunities to revise and use their local criteria during full-staff moderation meetings, that teachers developed school-wide understandings of these expectations. The findings that have been presented here suggest that such opportunities are more likely to occur if a commonly understood, school-wide moderation plan is developed. Likewise, agreement must be reached about who has responsibility for leading a school’s moderation processes.

Factors affecting the dependability of teacher judgements

This section reports on those occasions when teachers used their involvement in moderation to negotiate meaning about factors that could affect the dependability of their judgements of student achievement. As well as detailing instances in which teachers arrived at understandings that would allow them to strengthen the dependability of their judgements, it also describes some of the challenges that arose as
teachers engaged in this complex assessment activity. Two of the scenarios that are presented here explain how taking part in moderation provided teachers at Greenville School with opportunities to inquire into their judgement-making processes. The first of these is linked with one teacher’s experience of an inter-syndicate judgement-making session (see Appendix F, Meeting 4). It explores how Susan, the leader of this school’s middle syndicate, used her experience of this meeting to broaden her understanding of judgement making.

This session was the only occasion when teachers at Greenville School were observed moderating their judgements of student work. During this meeting, teachers from the middle and senior syndicates met to discuss their judgements of a common assessment task. Because this school’s special education needs co-ordinator (SENCO), Anna, worked closely with a number of students at this level, she also took part in this meeting. Prior to this session, the teachers in the middle and senior syndicates had worked collaboratively to conceptualise a series of school-specific assessment criteria. Further details about the steps involved in the development of these criteria, which they referred to as their Must, Should, Could charts, were provided earlier in this chapter. To prepare for this session, the teachers asked their students to write about a recent visit to a portable planetarium. After conducting this common assessment task, and before attending the judgement-making meeting, the teachers used their school-specific assessment criteria to evaluate a number of their own students’ responses.

During the judgement-making session, the teachers paired up with a colleague and swapped writing samples. After this, and with no further discussion about how they should go about participating in the collaborative judgement-making process, the teachers began to discuss their judgements of student writing. Early in this process, it became apparent that Susan had developed an elaborate system for masking both the identity of her students and her provisional judgement of each script. Perhaps because of this system, which one participant referred to as “the enigma code” (OBS: GS, 4, IS, P. 15, L. 7), the teachers focused primarily on comparing their judgements of student achievement. The following exchange exemplifies the discussions that teachers engaged in as they participated in this moderation session:
Anna: I’ve put “at” [the standard] for this one.

Susan: The other one?

Anna: Yeah.

Susan: And that’s where I’ve put him. I’ve put him “at.”

Anna: And there’s quite a few spelling errors too.

Susan: And I would have given him a chance too. I would have actually given him “two proficient” too but we are not doing that anymore.

Anna: No. He’s “at.” I’d say, this one.

Susan: He’s good, he’s solid, he’s not a way behind, he’s not a way ahead.

(OBS: GS, 4, IS, P. 16, L. 34 – P. 17, L. 4)

Within this discussion, the teachers repeatedly stated the judgement that they had each arrived at about this student’s writing (lines 1, 4, 9, & 10–11). Yet no reference was made to the assessment criteria that they had agreed to use for this purpose. Perhaps because of this, the teachers’ reasons for attaining their judgements were neither fully explicated nor rigorously examined. Although levels of inter-rater consistency were not formally recorded during this meeting, the teachers afforded priority to gauging the level of agreement that existed between their judgements (lines 1–4). This suggests that they understood the purpose of this procedure was to establish the reliability, as opposed to the validity, of their judgements.

As the following exchange reveals, this focus on the reliability of their judgement making was evident in other conversations that the teachers engaged in.

Susan: That’s exactly what I’ve put . . . I’ve put her “above” the National Standard.

Jack: Yes, she’s strong.

Susan: Yeah.

Jack: Great structure, very good structure.

Susan: And that makes sense, doesn’t it . . . Jack and I know what we’re doing. That’s interesting Jack, very interesting.

(OBS: GS, 4, IS, P. 15, L. 44 – P. 16, L. 12)
Here, the experience of reaching agreement with her colleague Jack seemed to bolster Susan’s confidence in their judgement-making procedure (lines 6–8). It might be inferred from this that Susan felt that reaching agreement in this way had validated their judgement-making process. Yet as was the case in the previous example, Susan and Jack did not seek to identify why they had reached the same conclusion about this student’s writing. This is problematic because appraising the validity of a judgement involves evaluating the quality of the inferences and interpretations on which that judgement was based. In each of these examples, the teachers did not subject their judgements to such an evaluative process.

There were indications, however, that the teachers at this school might pay greater attention to validity during subsequent judgement-making sessions. This possibility was linked with a conversation that Susan later engaged in at a senior staff meeting. During this meeting (which was not observed), Susan reportedly talked with the deputy principal, Christine, about the moderation session that she had recently facilitated. Susan later recounted this conversation to her middle- and senior-syndicate colleagues during an inter-syndicate moderation meeting (see Appendix F, Meeting 6). As the following statement reveals, Susan’s conversation with Christine prompted her to reflect on the judgement-making process that she and her colleagues had engaged in. She explained,

1 She [Christine: the deputy principal] brought this up at [the]
2 senior staff [meeting]. She was just saying, it would be really
3 good, rather than doing what I did . . . (I levelled it, and it was
4 a bit like a Lotto, who is going to win, who is going to guess
5 what I guessed?); that they [the teachers] just come . . . and
6 you have a photocopy of the same piece [of student writing],
7 and you have a good open discussion.

(OBS: GS, 6, IS, P. 56, L. 29–32)

Within this statement, Susan critiqued the judgement-making approach that she and her colleagues had previously employed, likening it to game playing and guesswork (lines 3–5). In this way, she appeared to acknowledge the limitations of a judgement-making process that focused primarily on gauging inter-rater consistency or reliability. In contrast with this, the alternative that she proposed (lines 5–7) seemed to recognise the importance of creating opportunities for teachers to discuss why and how they had arrived at their judgements of student achievement. Such an approach to judgement
making should enable teachers to pay greater attention to the validity of their judgements. These data suggest that involvement in moderation enabled Susan to conceptualise, or reify, a more dependable judgement-making procedure.

For Susan, the opportunity to observe another syndicate’s judgement-making procedure provided her with additional insights into this assessment activity. This experience allowed Susan and one of her middle-syndicate colleagues to identify the benefits associated with arriving at a commonly agreed judgement-making procedure. This is important because inconsistencies in the ways in which assessors arrive at judgements of student achievement can threaten the validity of assessment information (Crooks et al., 1996). In the scenario that is described here, the teachers involved learned how differing approaches to judgement making affect the inferences that teachers arrive at about student achievement. This learning was the product of a visit that Susan made to a junior-syndicate moderation meeting. As previously explained, the teachers in this syndicate chose not to adopt the Must, Should, Could system that their colleagues in the middle and senior syndicates had established. Instead, they developed a system within which the writing expectations associated with level 1 of the New Zealand Curriculum (Ministry of Education, 2007) were divided into 5 sublevels. They referred to these as levels 1a–e.

As Susan observed her junior-syndicate colleagues using these levels to describe their students’ writing, the teachers identified a discrepancy in the way in which they were applying these levels. Susan later shared this experience with her middle-syndicate colleagues. Here Susan and Phillip discussed the insights that this observation afforded them:

```
1 Susan: Right at the very start [of the junior-syndicate meeting] the question came up, well, you’ve got them 2, no
2 say [level] 1b, does that mean they are [working] there, or
3 they’ve achieved that? So that came up right at the start. And
4 you’ve just got to be so clear on all these things. Some of
5 them [the teachers] saw that that’s where they’re [the
6 students] working on. Some of them saw it as, they’ve
7 achieved that [level]. . . . So there’s so many things that we’ve
8 really got to nail down.
9 Phillip: So even when you’ve put those really quite tight
10 assessment criteria in place, there’s still a lot of room for
```
During this conversation, Susan described the inconsistency that the teachers in the junior syndicate had identified in their judgement-making processes (lines 2–4 & 5–8). She explained that some teachers were describing a student’s achievement as being at a particular level if the student was working within this level (lines 2–3 & 4–7), while others were only using this description if a student had achieved all the relevant criteria (lines 4 & 7–8). This experience seemed to broaden Susan’s appreciation of the factors that teachers must reach agreement about to ensure the dependability of their judgement-making processes (lines 5, 8–9, & 13–15). Susan’s description of these events illustrates how participating in the junior syndicate’s moderation meeting enabled her to negotiate, or reify, new understandings about judgement making.

Likewise, talking with Susan enabled Phillip to expand his conception of judgement making. Specifically, it prompted him to conclude that agreement must be reached about more than just the content of assessment criteria (lines 10–12). Because this experience alerted both of these teachers to an issue that could affect the quality of their judgements of student achievement, it provided them with an insight that they could use to strengthen their own judgement-making processes. In this regard, Susan’s comment about the things that she and her colleagues still needed to “nail down” (lines 8–9) indicates that this was an issue that she felt required further attention.

Within this section, the examples that have been presented so far have described instances in which involvement in moderation enabled teachers to identify ways that they could strengthen their judgement-making processes. There was, however, at least one occasion when involvement in moderation resulted in teachers arriving at an understanding that threatened the dependability of their judgements of student achievement. This episode occurred at Riverside School during a moderation cycle (see Appendix G, Meetings 2–4) that involved the use of three National Education Monitoring Project (NEMP) functional writing tasks (Crooks et al., 2007). NEMP ran in New Zealand from 1995 to 2010 (Educational Assessment Research Unit, 2010). Throughout this period it provided comprehensive, nationally representative assessment information about the knowledge, skills, and abilities of students at two levels: year 4
(ages 8–9) and year 8 (ages 12–13; Ministry of Education, 2011c). Although NEMP assessment tasks were designed for students at these levels, many of these tasks have been deemed appropriate for students of a much wider range of ages and curriculum levels (Gilmore, 2002). It is important to note, however, that these tasks have not been aligned with either the revised New Zealand Curriculum (Ministry of Education, 2007) or the National Standards (Ministry of Education, 2009c).

At Riverside School, the principal took responsibility for planning her school’s moderation activities. In this capacity, she sourced the tasks that the teachers used for moderation purposes and developed a timetable for their moderation activity. Despite the central role that the principal played in this regard, she did not attend the teachers’ syndicate-level moderation sessions. Instead, the leader of each syndicate facilitated these sessions. It was at one of these meetings (see Appendix G, Meeting 3) that teachers in the senior syndicate developed a process that they believed would allow them to convert the scores that students had attained on each NEMP task into a National Standards-type rating (e.g., well below, below, at, and above). To do this, the teachers decided to link these ratings with the total scores associated with each NEMP task. In the statement that follows, Stephen used this approach to propose a rationale for using each student’s total score on the task “Torch” for this purpose (Crooks et al., 2007, pp. 38–39; see Appendix P). He suggested, “So, ‘above’ would be [a total score of] 6 to 7, ‘at’ 4 to 5. Approaching or ‘below’ would be 2 to 3. And ‘well below’ is 1” (OBS: RS, 3, FS, P. 16, L. 19–20). The teachers decided that they would use this approach to generate a National Standards-type rating for each student’s response to each NEMP task. They agreed that this information would then be used to generate an overall teacher judgement to describe each student’s achievement in writing.

The likelihood that assessment information will be misinterpreted is greatest when those making inferences from the data were not involved in the task’s design (Crooks et al., 1996). Although Riverside School’s principal was aware that the scores associated with these NEMP tasks were not aligned with the National Standards, she was not present at the senior syndicate’s judgement-making meeting. In keeping with her understanding of the NEMP resources, her intended goal for this moderation cycle had been to provide teachers with opportunities to develop consistency in administering and marking common assessment tasks. When she later became aware of the way in which the teachers in the senior syndicate had used the NEMP data she commented,
But the purpose initially wasn’t to use it [these data] as a judgement against National Standards and a level of ability. . . . Because we weren’t . . . assessing for an actual grade of achievement. We were assessing for moderation. . . . This was really around our ability to moderate consistently, make judgements that are on a par.

(OBS: RS, 4, FS, P. 2, L. 41–13)

This statement reveals the discrepancy that existed between the principal’s intention for this moderation cycle (lines 3–6) and the way in which the members of the senior syndicate understood its function (lines 1–3). This discrepancy led the teachers in the senior syndicate to develop an inappropriate method for interpreting the scores that their students had attained on each of the NEMP tasks. This provides an example of the way in which mutually engaging in moderation processes can lead teachers to negotiate unintended, or inappropriate, meanings about assessment.

The point at which this moderation process took place within Riverside School’s assessment and reporting cycle could explain why the teachers in the senior syndicate misunderstood the principal’s intentions. Specifically, this process occurred while the teachers were preparing to compose written reports of student progress for parents and whānau. These reports require teachers to communicate student progress in relation to the National Standards (Ministry of Education, 2010d). This may explain why these teachers misinterpreted the principal’s instructions and deviated from the intended process. Unfortunately, this deviation led the teachers in the senior syndicate to negotiate an inappropriate method for interpreting student assessment information. This threatened the dependability of their judgements of student achievement. Arguably, this situation might have been avoided if the principal, who was leading this school’s moderation processes, had been more closely involved in these activities. Given her understanding of the role and function of the NEMP assessment tasks, it is likely that she would have ensured that the data from these assessments were used appropriately.

Although involvement in moderation led the teachers from Riverside School’s senior syndicate to engage in some unintended assessment activities, there were indications that this episode may have generated at least one positive learning opportunity. Specifically, this episode appeared to broaden one teacher’s understanding of the level of care that must be taken when selecting assessment tasks. Further details about this learning opportunity are provided in the next section. It was also evident that
a number of teachers at Greenville School used their participation in moderation to arrive at understandings that would enable them to strengthen the dependability of their judgements of student achievement. This suggests that moderation has the potential to provide teachers with opportunities to inquire into and improve their judgement-making processes.

**Factors to consider when selecting, designing, and administering assessment tasks**

The following section describes how involvement in moderation shaped teachers’ understandings of factors that affect the dependability of the assessment information that they gather. In this way, it shows how engaging in moderation alerted teachers to things that require attention when selecting, designing, and administering assessment tasks. With one exception, the findings that are presented here describe understandings that teachers negotiated as they participated in judgement-making sessions. These learning opportunities did not prompt teachers to adjust the inferences that they made about students’ achievement in relation to these assessment tasks. Instead, teachers negotiated ways to mitigate or address these issues in future assessment cycles.

At Central Heights School, engaging in the development of school-specific Success Criteria enabled one group of teachers to strengthen the design of an assessment task. This experience alerted them to the importance of ensuring that assessment tasks are well aligned with the learning domain that they are designed to assess. This learning opportunity took place during a meeting at which the teachers in this school’s senior syndicate worked collaboratively to draft Success Criteria for personal experience writing (see Appendix E, Meeting 1). As previously explained, the teachers at Central Heights School made extensive use of the centrally developed writing standards (Ministry of Education, 2009c, 2010b) as they constructed their local Success Criteria. During this particular meeting, the teachers made 47 references to these resources. Because these documents provide little information about the language features associated with specific writing purposes, or genres, these teachers were also observed making use of a number of other textual resources. During the meeting that is described here, the teachers made 24 references to the New Zealand Curriculum English Exemplars (Ministry of Education, 2003b). This resource contains goals, progress indicators, and exemplars that describe and illustrate the features of personal experience writing. Within this session, the teachers made use of this resource to check that their criteria included the relevant language features. Engaging in this cross-checking process provided them with confidence that their Success Criteria described the attributes of
personal experience writing. Evidence of this can be seen in the following statement, in which Kerry appraised her syndicate’s use of this resource:

And that [our Success Criteria] covers all of those things: “impact, voice, ideas, structure, sentences, vocab, language features” [reading from the headings listed in the New Zealand Curriculum English Exemplars progress indicator matrix for personal experience writing]. . . . We’ve got that sorted.

(OBS: CHS, 1, FS, P. 16, L. 5–17)

For these teachers, the opportunity to engage in a rigorous criteria-development process provided them with an important foundation for the next phase of their moderation work. This involved developing the common assessment task that they intended to use later in the teaching and learning cycle.

Almost as soon as the teachers embarked upon this activity, the conversation turned to using a forthcoming school trip to a local museum as the stimulus for this common assessment task. As the excerpt below reveals, however, the teachers quickly identified that the proposed task was not sufficiently aligned with the planned writing focus:

Kerry: It will be hard to get impact out of the Early Settlers [Museum]. . . . ’Cause that’s looking very like personal recount . . . as opposed to an experience which in some ways tantalises you. Do you know what I mean? To me there’s a difference.

Chris: Yeah.

Margaret: Yeah.

Chris: You’re retelling what you did as opposed to

William: It’s just a personal recount.

Kerry: You’re not going to get, yeah. I think we should do

Chris: Favourite holiday thing or something.

Kerry: Or my special place.

(OBS: CHS, 1, FS, P. 37, L. 46 – P. 38, L. 25)

This discussion reveals the teachers’ collective realisation (lines 2–9) that the content of the proposed assessment task was more closely aligned with personal recount writing (lines 3, 8, & 9) than the learning domain that they were intending to assess. This
realisation occurred because the experience of developing, or reifying, Success Criteria had allowed these teachers to negotiate a shared understanding of the learning domain associated with personal experience writing. It was this shared understanding that enabled Chris and Kerry to propose more appropriate topics (lines 11–12) for this assessment activity. This led to the development of a task that provided their students with the opportunity to demonstrate their ability to craft a piece of personal experience writing.

For a validity argument to encompass content considerations, it is necessary to evaluate how well an assessment task represents or samples the intended learning domain (Darr, 2005b). For these teachers, involvement in the development of local assessment criteria provided them with a sufficiently robust understanding of the learning domain to engage in this process. This example demonstrates how involvement in moderation can broaden teachers’ understandings of factors that affect the dependability of assessment information. It also illustrates how participating in moderation can enable teachers to strengthen the dependability of the assessment information they gather.

At each of the participating schools, engaging in moderation processes alerted teachers to the ways in which the design of an assessment task can shape the opportunities that students are afforded to display what they know and can do. This is important because the opportunity students are provided to demonstrate their knowledge and ability in the specified area is a key feature of the validity of any assessment (Harlen, 1994). At Central Heights School, teachers negotiated meaning about this issue during a full-staff judgement-making session (see Appendix E, Meeting 8). Although the teachers at this school typically used their locally developed Success Criteria to inform their conversations about student writing, this judgement-making meeting was an exception. Because this meeting had been scheduled to disseminate information about recent updates to the standardised assessment tool, e-asTTle writing (Ministry of Education, 2012a), the teachers instead used the criteria associated with the e-asTTle tool to inform their conversations about student writing. To do this they used the e-asTTle marking rubric (Ministry of Education, 2012b) to discuss and reach agreement about their judgements of an e-asTTle student exemplar entitled “Dogs at the Beach” (Evaluation Associates, 2012, p. 58). The teachers had been provided with the instructions associated with this task and these included the following directions:
Imagine you are at the beach watching these dogs [pictured]. Write to describe that moment in time: Give as many details as you can. Describe:

- What you see
- What you smell
- What you hear
- What you are thinking as you watch the dogs

(Evaluation Associates, 2012, p. 58)

Having read these instructions and reflected on the content of the student’s exemplar, one of the teachers, Chris, voiced a concern about the task. His comment prompted the following discussion:

1 Chris: I just think that it’s sort of interesting, you know
2 when they’ve [the e-asTTle developers] got a task like that
3 [Dogs at the Beach], like do they have a choice . . . for the
4 kids to write about? I’m thinking, like if you’re a kid who’s
5 got dogs and been to the beach then you’d be hugely
6 advantaged. . . .
7 Kerry: And you’re going to see that in the writing aren’t
8 you?
9 Chris: That’s what I mean.
10 Kerry: The ideas are going to be much, much richer.
(OBS: CHS, 8, SW, P. 3, L. 15–36)

Through their participation in this discussion these teachers negotiated meaning about the effect that prior experiences (lines 4–6), unrelated to the intended learning domain, could have on the opportunities that children were afforded to demonstrate their knowledge and skills through a given assessment task (lines 6–10). They concluded that those children who had previously experienced taking dogs to the beach were better positioned to write an effective response to this task.

This conversation also prompted those present at this meeting to consider the adverse effects that either a lack of prior experience, or a negative prior experience, could have on a child’s ability to respond to an assessment task. This opportunity
occurred when the principal of Central Heights School asked Chris for further clarification about the issue that he had raised:

1 Principal: So . . . I want to know whether I’ve got what you mean by this. It’s actually when you select the task for the group of children, you are selecting in a way that will be meaningful for them . . . Is that what you mean?
2 Chris: Yeah. Some children would find that [the “Dogs at the Beach” task] a doddle. You know like my sister takes her dogs with her kids down to the beach all the time, so they’ve obviously got a lot of knowledge to go with that.
3 Principal: Yeah.
4 Chris: Whereas, some poor little kid who didn’t get out.
5 Margaret: Who never went to the beach and didn’t have dogs, didn’t like dogs.

(OBS: CHS, 8, SW, P. 4, L. 15–37)

Within this exchange, the principal’s opening statement illustrates how moderation conversations provided participants with opportunities to inquire into and negotiate new meaning about complex assessment issues. In this case, her inquiry prompted Chris and Margaret to consider how students who had had either limited (lines 10–11) or negative (line 12) task-related experiences would perform on such an assessment. Issues like this deserve careful consideration because they can have an effect on the ways in which students respond to an assessment task.

For the teachers at Central Heights School, having the opportunity to discuss how different students might perceive and therefore respond to the “Dogs at the Beach” task (Evaluation Associates, 2012, p. 58) brought the issue of task design to the fore. Having deepened their understanding of those factors that shape the extent to which an assessment task enables a student to demonstrate their knowledge, skills, and abilities, the teachers then sought to apply these newly garnered insights. Here, they discussed how to use these principles to inform the development of a school-wide common assessment:

1 Margaret: Wouldn’t it be interesting to use that same task [the “Dogs at the beach” task] but with a photograph of a shared experience. For example if we
Within this exchange, the adjustments that were proposed to the “Dogs at the Beach”-style task (lines 2–4 & 7) provide an indication of the meaning that these participants had gleaned from their involvement in this session. The suggestion of using a shared experience to motivate a common writing assessment indicates that these teachers had arrived at a heightened understanding of the effect that prior experiences (unrelated to the intended learning domain) can have on the opportunity an assessment affords children to demonstrate their knowledge and skills.

Yet as the teachers at Greenville School learned, using a common experience as a motivation for a writing assessment does not ensure that students will have the opportunity to demonstrate their skills and abilities. This meaning was negotiated as teachers in Greenville School’s middle and senior syndicates discussed student responses to a common writing assessment (see Appendix F, Meeting 4). For this task, students had been asked to write a recount describing their recent experience of a local museum’s portable planetarium. As the teachers participated in this judgement-making session, the opportunity to compare and discuss their judgements of student writing prompted a conversation about the assessment task:

1 Richard: But it wasn’t their best writing because it [the planetarium] wasn’t the most exciting thing.
2 Jack: And that’s why I said a lot of them were down.
3 Anna: No, it [the planetarium] wasn’t [very exciting]. So that’s when your OTJ [overall teacher judgement] is going to come in because it [the planetarium] didn’t actually inspire them.
4 Jessica: No it didn’t.
5 Anna: I know how they [the students] feel, because I left feeling, so, yeah.
6 Jack: A rocket would be [a] good [task stimulus].

(OBS: GS, 4, IS, P. 20, L. 21–41)
Here, the teachers negotiated meaning about how an apparent lack of interest in the assessment task (lines 1–2, 4, 6–7, 8, & 9–10) had adversely affected student performance (lines 1 & 3).

As well as alerting these teachers to the effect that a poorly perceived assessment task can have on student performance, this discussion also provided them with a forum for inquiring into and addressing the task issue that they had identified. The teachers used this conversation to propose a number of methods for increasing the opportunities that students were afforded to demonstrate their knowledge and ability. One example of this can be seen in Anna’s “OTJ” (overall teacher judgement) comment (line 5). Within the National Standards context, teachers are expected to arrive at an OTJ of each student’s achievement in writing (Ministry of Education, 2009c). To make these judgements, teachers must collect and analyse a range of pieces of assessment evidence for each student. They are then expected to use this evidence to arrive at an overall judgement about each student’s progress and achievement in relation to the appropriate standard. Given this, Anna’s comment served as a reminder to all present that an overall judgement of a student’s knowledge and ability should not be reached on the basis of a single assessment. Her statement reflects the widely held view that increasing the number of information sources, or the size of the sample, on which a judgement is based also increases its dependability (Crooks et al., 1996; J. Smith, 2003). This further exemplifies the ways in which involvement in moderation drew teachers’ attention to the complexities associated with the collection of student assessment information. As well as alerting these teachers to an assessment design issue, this experience also prompted them to think about the importance of collecting a broad range of assessment information.

Participating in the moderation of a series of school-wide common assessment tasks led at least one teacher at Riverside School to draw an important conclusion about task selection. During this moderation cycle (see Appendix G, Meetings 2–4), the teachers at this school administered three NEMP functional writing tasks (Crooks et al., 2007) to their students. These assessment tasks, each of which required students to write a set of instructions, were not aligned with the New Zealand Curriculum (Ministry of Education, 2007). As explained earlier, the teachers in Riverside School’s senior syndicate misunderstood the purpose of this moderation exercise and used the resulting assessment information in unintended ways. Yet because this misunderstanding was later identified, this experience provided one member of the senior syndicate with a
valuable learning opportunity. This teacher, Kim, who taught a class of year 7 and 8 students, later talked about what she had learned from this experience:

1. Not every writing task will lend itself to actually show[ing] that the child can do that at level 4. . . . It’s just that not every task is going to really be curriculum level based. . . .
2. I guess the [NEMP] writing [tasks] highlighted that aspect as well, that not all writing tasks are going to meet what you [the teacher] are looking for.

Within this statement, Kim alluded to the NEMP tasks that she and her colleagues had used to collect information about student achievement (lines 3–5). Although these tasks were developed for pupils in both year 4 and year 8, they were not designed to measure whether students were meeting the writing demands that are specified at level 4 of the New Zealand Curriculum (Ministry of Education, 2007). This is the level at which New Zealand’s reading and writing standards state that year 8 students should be working (Ministry of Education, 2009c).

Because Kim taught year 8 students, she was looking for evidence that their writing fulfilled the level 4 expectations. She noted, however, that the NEMP tasks (Crooks et al., 2007) did not appear to provide her students with the opportunity to produce writing that demonstrated this level of proficiency (lines 1–2 & 4–6). Because Kim’s students had not been provided with the opportunity to demonstrate the extent of their knowledge and ability, she did not have access to the information that she required in order to make valid judgements of their achievement. Kim’s statement demonstrates how involvement in moderation heightened her understanding of the level of care that teachers must take when selecting the assessment tasks (lines 2–6) that they will later use to inform their judgements of student achievement.

Taking part in moderation also enabled two teachers from Central Heights School’s senior syndicate to learn how differences in assessment conditions can affect student performance. This learning opportunity was the product of a paired judgement-making session (see Appendix E, Meeting 10). Within this meeting, Kerry and Chris discussed their judgements of student responses to a character-writing assessment. For this assessment, the teachers in the senior syndicate had sought to develop and administer a common assessment task. In this regard, they had agreed that their students would all write a character description of a competitor at the ancient Olympics. As the
following exchange reveals, however, these teachers gained some additional insights into task design and administration as they discussed their judgements of student writing:

1. Kerry: And as you can see [from their writing], their [task] instructions were a bit different . . . between yours and mine. Which is interesting, because we have tried really hard on this task.
2. Chris: Yeah, to make it the same.
3. Kerry: to try and get the same character studies task. But they’re [the students are] very specific about the instructions that they follow . . . My kids have done exactly what they were asked and your kids
4. Chris: Yeah, they’re the same.

(OBS: CHS, 10, FS, P. 29, L. 8–42)

Although these teachers had intended to devise and administer a common assessment task (lines 3–6), the opportunity to discuss their students’ writing (line 1) alerted them to the fact that they may have introduced the task differently to their respective classes (lines 1–2). Because this collaborative judgement-making session had enabled each teacher to engage closely with the writing of their colleague’s students, they could see how these differing instructions had affected the students’ writing. This heightened their collective understanding of the effect that specific task instructions have on the work that students produce (lines 6–10).

Differences in the task instructions that are provided to students can pose a threat to validity because the conditions or circumstances of assessment affect the opportunity that students are afforded to demonstrate their skills and abilities (Crooks et al., 1996). Having identified the ways in which a difference in task administration had affected the student writing that she and her colleague had discussed, Kerry reflected further on this. She made the following comment in a post-moderation interview:

1. Kerry: And now I can see that there’s probably more moderating about how the teacher works the task before the kids have worked the task. Yeah, yeah, it’s quite interesting.
2. Esther: Just even in that conversation that you were having with Chris about the different
Kerry: Yeah, the different ways we had set the task up. . . .

It shows how complicated assessment is.

This exchange illustrates how participating in moderation expanded Kerry’s understanding, or reification, of task administration (lines 1–3). Although Kerry thought that she and her colleagues had agreed on a common assessment task, involvement in moderation alerted her to differences in the ways that she and Chris had introduced this task to their students (lines 6–7). Likewise, her attention was also drawn to the effect that these differences had on the resulting student writing. This led her to conclude that there were more factors to consider when designing a common assessment task than she had previously realised (lines 1–3).

The findings that have been presented within this section illustrate how engaging in moderation can enable teachers to negotiate meaning about factors that affect the dependability of the assessment information that they collect. In each of the scenarios that have been depicted, there were indications that the teachers concerned intended to use the insights they had gained to improve their processes for designing, selecting, and administering assessment tasks. This suggests that participating in moderation has the capacity to strengthen the assessment capability of teachers in ways that should lead to improvements in the dependability of the assessment information that they gather.

Chapter conclusion

This chapter has shown that involvement in moderation enabled teachers to negotiate new understandings about factors that affect the dependability of assessment information. It has argued that participating in moderation provided teachers with opportunities to develop, or reify, shared understandings of their local assessment criteria. Likewise, it identified that teachers used their involvement in these processes to garner insights into those factors that affect the dependability of qualitative judgements of student achievement. Finally, this chapter has shown that participating in moderation enabled teachers to improve their ability to design, select, and administer assessment tasks. Throughout this chapter, attention has also been paid to the ways in which the leadership of each school’s moderation processes shaped the learning opportunities that teachers were afforded. The findings that have been presented here indicate that
teachers benefited optimally when they participated in carefully planned, coherently led moderation processes.
Chapter 6. Moderation and assessment identity: Teachers’ accounts of how involvement in moderation shaped their assessment capability

This chapter uses Wenger’s (1998) conception of identity, or “learning as becoming” (p. 5), to explore how involvement in moderation shaped teachers’ perceptions of their assessment capability. To do this, it draws on teachers’ accounts of how and what they learned about assessment through participating in moderation.\textsuperscript{11} The chapter begins by examining teachers’ descriptions of how taking part in moderation informed their understandings of both the centrally developed writing standards (Ministry of Education, 2009c, 2010b) and their own school’s writing expectations. The next section, which is divided into two parts, explores teachers’ accounts of how involvement in moderation affected their ability to both arrive at judgements of student writing and use student assessment information to inform teaching decisions. The findings that are presented within this section provide insights into the ways in which teachers perceived that participating in moderation informed their ability to engage in a process that the New Zealand Ministry of Education (2007) has called “learning inquiry” (p. 35). Throughout this chapter, attention is drawn to the reports that teachers provided about the ways in which developing and using school-specific assessment criteria improved their assessment capability. It is argued that those teachers whose experiences of moderation included taking part in these activities were the most likely to report that moderation had strengthened their assessment capability.

Teachers’ views about how involvement in moderation shaped their understandings of writing standards and achievement expectations

This section examines teachers’ accounts of how taking part in moderation shaped their understandings of both the centrally developed writing standards (Ministry of Education, 2009c, 2010b) and their own school’s writing expectations. It shows that those teachers whose experiences of moderation included using the centrally developed writing standards to co-construct school-specific assessment criteria were the most likely to talk about how involvement in moderation had enabled them to clarify their understandings

\textsuperscript{11} Because most of these accounts were provided during interviews, the voices of the teachers in each school’s focus syndicate can be heard the most clearly in this chapter. The perceptions of other teachers are, however, reflected in the questionnaire data that are presented in Figures 2–4.
of writing standards and expectations. This is important because research (Wyatt-Smith et al., 2010) has suggested that teachers require opportunities to discuss and clarify their interpretations of such descriptions of achievement to enable them to arrive at shared understandings of them. These common understandings are necessary because teachers draw on them to inform the judgements of student achievement at which they arrive (Klenowski & Wyatt-Smith, 2010a, 2014; Maxwell, 2010). The findings that are presented here provide insights into the ways in which teachers perceived that participating in moderation informed their understandings of both the centrally developed writing standards (Ministry of Education, 2009c, 2010b) and their own school’s writing expectations. This allows conclusions to be drawn about the ways in which teachers felt that taking part in moderation contributed to this aspect of their assessment capability.

A community defines what it means to be a competent member through its practice (Wenger, 1998). Within any community, the ability to make use of the shared repertoire of that practice is a marker of competent membership. A community’s shared repertoire includes the tangible resources, intangible understandings, and commonly agreed systems and procedures that its members have created and claimed to enable them to sustain their mutual engagement in action. Because a member’s ability to interpret and make use of that community’s shared repertoire is predicated on having a personal history of engaging with and shaping that repertoire, this dimension of competence becomes a dimension of identity. For this reason, being able to negotiate or make use of the repertoire of a practice is both a marker of competence and a dimension of identity. Since competence and identity are so closely interrelated, these terms are used interchangeably within this chapter.

The centrally developed writing standards (Ministry of Education, 2009c, 2010b) featured in the shared repertoire of each of the participating schools. At Riverside School, for example, teachers were observed using the National Standards (Ministry of Education, 2009c) to inform their judgements of student writing (see Appendix G, Meeting 1). Likewise, teachers from both Central Heights and Greenville schools made systematic use of the Literacy Learning Progressions (Ministry of Education, 2010b) and the National Standards during the development of their local assessment criteria. At both of these schools, these locally developed resources became part of the shared repertoires of their respective practices. At Central Heights School, these resources were referred to as their Success Criteria (see Appendix B). Likewise, at Greenville School the child-friendly versions of these criteria were called Must, Should,
Could charts (see Appendix C), and the teachers’ versions were referred to as Tracking Sheets (see Appendix Q).

Figures 2–4 show that only teachers at the two schools whose moderation activities included the creation of local assessment criteria strongly agreed that involvement in moderation had enabled them to clarify their understandings of both the centrally developed writing standards (Ministry of Education, 2009c, 2010b) and their own school’s writing expectations.

**Figure 2. The level of agreement that teachers from Central Heights School reported in relation to the statements “Moderation is helping me to clarify my understanding of (a) the National Standards for writing; (b) the Literacy Learning Progressions for writing; and (c) my own school’s writing expectations.”**

**Figure 3. The level of agreement that teachers from Greenville School reported in relation to the statements “Moderation is helping me to clarify my understanding of (a) the National Standards for writing; (b) the Literacy Learning Progressions for writing; and (c) my own school’s writing expectations.”**
Teachers from both Central Heights and Greenville schools were also much more likely to provide verbal accounts of the ways in which their involvement in moderation had positively shaped their understandings of writing standards and expectations than were their peers at Riverside School. For example, Susan, a teacher from Greenville School, talked about how taking part in the criteria-development process had enabled her to clarify her understandings of the centrally developed writing standards. She explained, “Having gone through this [the development of our Must, Should, Could charts] . . . we are very aware now of exactly what’s in those documents [the centrally developed writing standards] and what the expectations are” (INT: GS, 1, FST, GRP, P. 12, L. 43–45).According to Susan, it was the process of using the centrally developed documents (Ministry of Education, 2009c, 2010b) during the creation of her school’s local assessment criteria that enabled her to strengthen her understanding of the official, centrally developed writing standards. For Susan, engaging in this aspect of her school’s moderation practice increased the confidence that she had in her ability to negotiate or make use of this aspect of their shared repertoire. This provides an indication of how the experience of developing local assessment criteria shaped Susan’s assessment identity.

Susan’s colleague, Phillip, articulated a similar sentiment. He explained,

1 I think the process of actually developing the Musts, Shoulds
2 and Coulds actually gives you a vague idea of where the
3 levels are, what [National] Standard is within each level. . . .
4 From my point of view that [developing the Must, Should,
5 Could charts] was really useful for making it very clear what

Figure 4. The level of agreement that teachers from Riverside School teachers reported in relation to the statements “Moderation is helping me to clarify my understanding of (a) the National Standards for writing; (b) the Literacy Learning Progressions for writing; and (c) my own school’s writing expectations.”
level I should be expecting from these kids.

Although Phillip spoke somewhat cautiously about the effect that involvement in moderation had had upon his understanding of the National Standards (Ministry of Education, 2009c; lines 1–3), he too linked his participation in the development of school-specific assessment criteria (lines 1–2) with having the opportunity to clarify his understandings of these centrally developed writing standards (lines 2–3). Moreover, Phillip’s statement also indicates that he perceived that taking part in the development of the Must, Should, Could charts had enabled him to clarify his understanding of his own school’s writing expectations (lines 3–6). Becoming a competent member of a community involves both developing the ability to recognise competence and being recognised by others as competent (Wenger, 1998). For Phillip, involvement in moderation strengthened his ability to recognise competence in two interrelated ways. Because taking part in criteria development increased the confidence that he had in his ability to recognise his students’ competence (lines 1–3 & 4–6), this experience allowed him to develop a clearer understanding of what it meant to be a competent assessor. This seemed to strengthen Phillip’s conception of his assessment identity.

A number of teachers from Central Heights School also talked about how their use of the centrally developed writing standards (Ministry of Education, 2009c, 2010b) had shaped the construction of their school’s Success Criteria. Their comments provide insights into the ways in which these teachers perceived that involvement in moderation informed their understandings of writing standards and expectations. For example, Chris commented that participating in the development of his school’s Success Criteria had required him to engage with and interpret the centrally developed writing standards:

1 You actually think, once you’ve got the [Success] Criteria, and especially we worded it a lot easier and it makes much more sense than what’s in there [the centrally developed writing standards]. Because there’s quite a lot of subjectivity with some of them in what they are asking you to do. . . .
2 Like you’re looking at it and thinking, what do they actually mean by that?

Within this statement, Chris referred to the interpretive process (lines 4–7) that he and his colleagues had taken part in as they used the centrally developed writing standards
to create their Success Criteria (lines 1–4). Here, Chris’s comment about his own school’s descriptions of achievement making more sense than those that were specified in the centrally developed writing standards (lines 2–4) indicates that he felt that his involvement in the production of his school’s Success Criteria had allowed him to clarify his understanding of these local expectations. At the same time, it might also be inferred that Chris recognised that he could not have achieved this clarity without first interpreting and deciphering the centrally developed writing standards. That is, Chris seemed to acknowledge that to arrive at his school’s more readily understood local Success Criteria (lines 2–3), he and his colleagues had first had to identify what the authors of the centrally developed writing standards had actually meant to convey (lines 6–7). Because Chris perceived that taking part in criteria development strengthened his ability to interpret and make use of each of these aspects of his community’s shared repertoire, it is argued that his involvement in these processes had a positive effect on his assessment identity.

Like Chris, Margaret also reported that Central Heights School’s own writing expectations were easier to understand and make use of than those specified in the centrally developed writing standards (Ministry of Education, 2009c, 2010b). She explained,

1 We used those, the [National] Standards and the Literacy [Learning] Progressions . . . we’ve used those to develop
2 these [our Success Criteria]. So I think in a way it’s,
3 simplified is the wrong word but . . . I can understand this
4 [our Success Criteria] and I can use this in my teaching.
(INT: CHS, 1, FST, GRP, P. 23, L. 44–46)

Here, Margaret’s account of her involvement in the development of local assessment criteria bears some resemblance to that of her colleague Chris. For example, like Chris, Margaret felt that using the centrally developed writing standards during the co-construction of her own school’s Success Criteria (lines 1–3) enabled her to clarify her understanding of these locally defined writing expectations (lines 3–5). It might also be inferred from Margaret’s comments (lines 4–5) that she, too, perceived that involvement in this process strengthened her understanding of the centrally developed writing standards. Perhaps most importantly, Margaret’s participation in the development of local assessment criteria provided her with the sense that she understood and could use these criteria in her teaching (lines 4–5). The confidence that
Margaret expressed in her ability to interpret and make use of her school’s shared repertoire provides an indication of how involvement in criteria development shaped her assessment identity. The next part of the chapter picks up on this idea and explores the ways in which teachers used their understandings of standards and assessment criteria to inform decisions about both student achievement and next teaching and learning steps. This allows further conclusions to be drawn about how the experience of moderation affected teachers’ assessment identities.

Moderation and learning inquiry: Teachers’ views about how moderation shaped their ability to respond to assessment information

The following section examines teachers’ accounts of what and how they learned about assessment through their involvement in moderation to determine how these experiences shaped their ability to conduct a learning inquiry. According to the New Zealand Ministry of Education (2007), a learning inquiry is comprised of two phases. During the first phase, teachers must ascertain “what happened as a result of the teaching” (p. 35). Within the second, teachers need to analyse the assessment information and identify the “implications for future teaching” (p. 35). To investigate teachers’ perceptions about the effect that involvement in moderation had upon their capacity to undertake each facet of a learning inquiry, this section is presented in two parts. It begins by exploring teachers’ accounts of how taking part in moderation shaped their ability to arrive at dependable judgements of student achievement. This provides insights into the extent to which teachers felt that involvement in moderation shaped their capacity to evaluate the effects of their teaching. The second part of this section examines the ways in which teachers thought that participating in moderation affected their ability to use assessment information to identify next teaching and learning steps. This provides an indication of the degree to which teachers felt that involvement in moderation shaped their capacity to undertake the second aspect of a learning inquiry. Because the findings that are presented in each of these sections reveal the ways in which teachers felt that participating in moderation affected their assessment competence, these data provide insights into how the teachers’ experiences of moderation informed their assessment identities.
Identifying what happened as a result of the teaching: Teachers’ views about how moderation shaped their judgement-making ability

Teachers at all three participating schools talked about how taking part in moderation had shaped their ability to arrive at judgements of student writing. Yet these accounts differed. The data that are presented in this section show that the teachers at Central Heights and Greenville schools, whose involvement in moderation included participating in the co-construction of local assessment criteria, were the most likely to describe how their experiences of moderation had strengthened their judgement-making ability. At both of these schools, teachers talked about how their involvement in criteria development had informed this dimension of their assessment competence. For example, Chris, from Central Heights School, explained,

1 So, I think that’s where, you know, if we’ve got something
2 that we’ve . . . identified together, in this stage [the criteria-
3 development stage], when you look through at a kid’s piece
4 of work, I mean, sure . . . if you were going to be really
5 picky there might be some things that you didn’t quite agree
6 perfectly on, but not a lot.

(INT: CHS, 2, FST, IND, P. 13, L. 48–51)

Here, Chris’s use of the phrase “something that we’ve both identified together” (lines 1–2) suggests that he equated the process of working collaboratively to construct Success Criteria (lines 1–3) with the emergence of common understandings of these descriptions of achievement. According to Chris, participating in the criteria-development process increased the likelihood that teachers would apply these criteria consistently when they used them to inform their judgements of student writing (lines 3–6). This indicates that he perceived that the shared understandings that he and his colleagues had arrived at during the criteria-development phase had a positive effect on their judgement-making ability. Wenger (1998) has argued that identity is formed through a “very complex interweaving of participative experience and reificative projections” (p. 151). Chris felt confident that he and his colleagues had the ability to make use of their Success Criteria during judgement making because they had participated in their development, or reification. In keeping with this view, he later explained, “So, I think it’s that level of agreement that you can have in there [the Success Criteria] makes it [judgement making] much easier” (INT: CHS, 2, FST, IND, P. 15, L. 5–6).
Like Chris, Phillip’s experience of moderation included contributing to the
development of school-specific assessment criteria. Phillip, a teacher from Greenville School, talked about the ways in which taking part in the construction of his school’s Must, Should, Could charts had shaped his ability to make and report on judgements of student achievement. He explained,

1 Well, it [involvement in moderation] makes you feel a lot
2 more confident talking to parents about where their child is,
3 rather than talking about a [National] Standard that you’re
4 not entirely sure about yourself. Which is, without having
5 done this [the development of the Must, Should, Could
6 charts] . . . that is kind of what you are left doing. And it
7 makes parents’ evening a lot less scary. . . . You don’t feel
8 like you are fluffing your way through it. You feel like you
9 actually know what you are talking about.

(INT: GS, 1, FST, GRP, P. 39, L. 12–33)

Within this statement, Phillip linked his involvement in the development of his school’s Must, Should, Could charts (lines 4–6) with having increased confidence in his ability to justify his judgements of student achievement to parents (lines 1–4 & 6–9). Because Phillip talked about how the experience of criteria development had strengthened his ability to display his assessment competence, this statement provides an indication of how taking part in moderation shaped Phillip’s assessment identity.

The teachers from both Central Heights and Greenville schools also talked in more general terms about the ways in which involvement in moderation had increased their confidence in their judgement-making ability. For example, Phillip’s colleague Susan explained, “Oh it’s [moderation has] made our judgements better. When I say a child is ‘below’ or ‘well below’ [the National Standard], I am pretty confident they are” (INT: GS, 1, FST, GRP, P. 41, L. 17–18). This was a sentiment that William, from Central Heights School, echoed when he was asked to reflect upon how taking part in moderation had affected the way that he made judgements of student writing. He replied, “I think I am more confident, more relaxed, about it [making judgements]. So then, I think I am going to be more realistic and honest and it’s [my judgements are] going to be more accurate” (INT: CHS, 1, FST, GRP, P. 12, L. 1–2). Likewise, William’s colleague Margaret talked about how involvement in moderation had shifted her attitude towards making judgements of student writing, or marking. She explained,
I tended to have, not a negative attitude but a “Oh god, I’ve got to do some marking.” And you know . . . it can be a bit tedious. I feel more confident about doing it [making judgements of student work] and I know what I am doing. So it’s [making judgements is] actually easier. . . So that it’s not that big “oh my goodness, I don’t want to mark all that writing.” It’s “I’ve got to mark the writing but I know what I’m doing.”

(INT: CHS, 1, FST, GRP, P. 12, L. 5–22)

Within this statement, Margaret linked her experiences of moderation with having opportunities to strengthen her ability to make judgements of student writing (lines 3–4 & 7–8). For Margaret, involvement in moderation was identity forming because she felt that these experiences had enabled her to get better at using her school’s judgement-making systems.

Conversations with the teachers from Riverside School, for whom participation in moderation had not included the development of school-specific assessment criteria, also provided some insights into how they perceived involvement in moderation had affected their judgement-making ability. Perhaps most notably, one teacher from this school, Kim, emphasised that moderation had not affected this aspect of her assessment competence. When asked to talk about whether taking part in moderation had shaped her judgement making, Kim was unequivocal. She stated, “I don’t think it [moderation] really has changed [it]” (INT: RS, 2, FST, IND, P. 5, L. 15). Likewise, Kim expressed a similar sentiment when she was asked to comment on how involvement in moderation had affected the way in which she reported her judgements of student achievement to parents. This sentiment was captured in the following exchange:

Kim: I don’t think it’s [moderation has] really enhanced that [reporting to parents]. It’s given you an outcome, but it really hasn’t changed the way that you would [report judgements].

Esther: Has it [moderation] made you feel differently about what you write down [in reports to parents]? Like, did you feel more or less confident or anything like that? . . .

Kim: No.

(INT: RS, 2, FST, IND, P. 9, L. 28–37)
In contrast with Kim, who insisted that involvement in moderation had not affected either her judgement-making practice (lines 1–3) or her ability to report to parents (line 7), Stephen held a different view. This was reflected in his response to a question about how he felt that moderation had shaped the way in which his school reported student achievement to their board of trustees:

1. I think that the teacher judgement is much, much better. . . .
2. And I think the insight or the snapshot [of student achievement data] that the board [of trustees] is getting is
3. much, much better than it used to be, through this
4. [moderation] process.

(INT: RS, 1, FST, IND, P. 13, L. 42–45)

This statement indicates that Stephen felt that involvement in moderation had strengthened the quality of the judgements that he and his colleagues arrived at (line 1).

Although Stephen was the only teacher at Riverside School to talk about the positive effect that taking part in moderation had had upon the quality of teachers’ judgements, the questionnaire data that are presented in Figure 5 indicate that his colleagues generally agreed with this perception. It is, however, also clear from this figure that higher proportions of teachers at Central Heights and Greenville schools expressed stronger levels of agreement with this sentiment than they did at Riverside School.

Figure 5. The level of agreement that teachers reported, by school, in relation to the statement “Moderation is making me feel more confident about making dependable judgements of student achievement in writing.” CHS = Central Heights School; GS = Greenville School; RS = Riverside School.
I argue that the difference in the levels of confidence that these teachers expressed, both qualitatively and quantitatively, about their judgement-making ability might be related to the quality of the understandings that they felt they had of the relevant standards and criteria. To apply stated standards consistently during judgement making, teachers require robust, common understandings of these descriptions of achievement (Klenowski & Wyatt-Smith, 2010a, 2014; Maxwell, 2010). Given that those teachers whose experiences of moderation included participating in the co-construction of local assessment criteria were the most likely to report that moderation had enabled them to clarify their understandings of writing standards and expectations, it logically follows that these teachers would also have the most confidence in their ability to use these standards to arrive at dependable judgements of student work.

Interestingly, the frequency with which teachers participated in judgement-making sessions did not seem to influence the likelihood that they would either report that taking part in moderation had improved their judgement making or strongly agree that these experiences had made them feel more confident about arriving at dependable judgements of student achievement.

Table 9

Number of judgement-making and criteria-development meetings that the members of each school’s focus syndicate are known to have participated in

<table>
<thead>
<tr>
<th>School</th>
<th>Number of judgement-making meetings</th>
<th>Number of criteria-development meetings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Heights</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>School</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenville School</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Riverside School</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

As shown in Table 9, teachers at Greenville School had fewer opportunities to participate in judgement-making sessions than those at Riverside School. Despite this, the teachers at Greenville School were much more likely than those at Riverside School to describe how taking part in moderation had strengthened this aspect of their assessment competence. Likewise, higher proportions of teachers at Greenville School than at Riverside School strongly agreed that these experiences had made them feel more confident about arriving at dependable judgements of student achievement.
The findings that have been presented in this section suggest that involvement in criteria development enabled teachers to clarify and arrive at shared understandings of standards. Teachers’ involvement in criteria development may therefore have helped strengthen their confidence in using criteria to arrive at dependable judgements of student writing. Because taking part in criteria development seemed to strengthen the confidence that teachers had in their ability to use their community’s judgement-making systems, it follows that involvement in criteria development may have contributed positively to teachers’ assessment identities.

**Identifying the implications for future teaching: Teachers’ views about how moderation shaped their ability to identify next teaching and learning steps**

This section examines the accounts that teachers provided about how involvement in moderation shaped their ability to use assessment information for formative purposes. These data complement earlier findings that showed how the teachers at Central Heights School used their experiences of moderation to establish a connection between judgement making and teaching as inquiry (see Chapter 4). Although Central Heights was the only school at which teachers systematically integrated the identification of next teaching and learning steps into their judgement-making processes, there were indications that the teachers from at least one other school were working towards this. It is argued here that those teachers whose experiences of moderation included participating in the co-construction of local assessment criteria were the most likely to use involvement in moderation for identifying next teaching and learning steps. Likewise, it is contended that these teachers were also the most likely to report that taking part in moderation had strengthened this aspect of their assessment competence.

At Greenville School, both Sophia and her colleague Susan described how using their locally developed assessment criteria helped them to identify next teaching and learning steps for their students. Susan talked about how utilising the Tracking Sheets (see Appendix Q) that she and her colleagues had adapted from their Must, Should, Could charts helped her to identify what she needed to focus on next in her teaching. She explained, “We got the Tracking Sheets and we took a few kids and we highlighted. And you suddenly see, oh I’ve been so busy trying to write, I haven’t been doing any editing” (INT: GS, 1, FST, GRP, P. 21, L. 32–34). For Susan, involvement in moderation helped her to “see” or identify what her next teaching focus needed to be. Likewise, Susan’s colleague Sophia reported having a similar experience. She commented,
You don’t feel, with the Musts, Shoulds and Coulds or with the Tracking Sheet, when you see the bits that are missing, you don’t feel so bad if they are missing full stops, just working on where full stops go.

(INT: GS, 1, FST, GRP, P. 22, L. 42–44)

Within this statement, Sophia described how utilising the resources that she and her colleagues had developed during their moderation practice (lines 1–2) increased the confidence that she had in her ability to use assessment information formatively (lines 2–4). Like Susan, Sophia found that engaging in this process helped her to see (line 2) what her students needed to be taught.

Phillip, another teacher from Greenville School, talked about why he felt that his school’s Must, Should, Could Charts would assist teachers with getting better at identifying the implications for future teaching. He explained,

Now there is a clear framework in place [the Must, Should, Could charts] . . . you are going to get more discussion of pedagogy. You know, how [do] you go about doing this? How [do] you address these issues? Well, this kid’s not doing this, this, and this. What more can you [the teacher] do? . . . So suddenly you’re getting into the real core of our job.

(INT: GS, 1, FST, GRP, P. 52, L. 17–30)

For Phillip, there was an important relationship between having a clear, readily understood set of criteria in place (lines 1–2) and the likelihood that teachers would actually use these criteria to inform their teaching decisions (lines 2–6). Phillip’s comment implied that it was because he and his colleagues understood and had ownership of their local criteria that they would be able to use them to inquire into and improve their teaching. For Wenger (1998), “ownership of meaning” (p. 200) and identity are closely linked. This is because having ownership of meaning involves having a degree of control over these meanings and therefore being able to utilise them.

The notion of ownership was something that Margaret, a teacher from Central Heights School, spoke about directly. This topic arose when she talked about why she and her colleagues felt more positive about using their locally developed Success Criteria than they did about utilising the National Standards (Ministry of Education, 2009c). She explained,
Because those [National] Standards were imposed on us, rather than consulted, it’s very difficult to feel any affinity with them. Whereas I feel really that we have got a lot of ownership of what we have done here because we have been involved in it [the development of the Success Criteria] and that’s why I think we feel so much happier about using something like this [the Success Criteria] than something that’s just been dumped on us and we’ve had no input.

(INT: CHS, 1, FST, GRP, P. 35, L. 33–37)

According to Margaret, it was because she and her colleagues had mutually engaged in the development of their Success Criteria (lines 4–5) that they felt they had “ownership” of this resource (lines 3–4). As Margaret explained, this sense of ownership assured teachers of the utility of their criteria (lines 6–7). Throughout this statement, Margaret talked collectively about the sense of ownership that she and her colleagues had in relation to their Success Criteria. This is something that Wenger (1998) has acknowledged. He explained that “ownership of meaning can be shared and it can have degrees. . . . It does not diminish from being shared” (p. 200). Because of the socially negotiated nature of meaning, establishing shared ownership can increase participation. This is an idea that Margaret touched upon when she talked about how happy she and her colleagues felt about using their locally developed Success Criteria (lines 6–7).

Margaret also talked specifically about how her school’s Success Criteria helped her to feel more confident about identifying next teaching and learning steps. When she was asked how taking part in moderation had affected her judgement-making practice, Margaret replied,

1 When you come to giving feedback on the writing, we’ve actually got specific things that we can give feedback on. So you don’t feel like you are trying to pull something out of the ether for every child. . . . You’ve actually got something concrete to say about what they can do and what their next steps are. And that’s really good.

(INT: CHS, 1, FST, GRP, P. 12, L. 31–40)

For Margaret, her school’s Success Criteria provided her with specific (lines 1–2), concrete (lines 4–5) things that she could give her students feedback about. Importantly,
this enabled Margaret to provide students with information about their next learning steps (lines 5–6). Because Margaret contrasted this with prior experiences of trying to provide feedback without the help of locally agreed criteria (lines 3–4), she drew attention to the effect that taking part in moderation had had upon this aspect of her assessment identity.

Chris, another teacher from Central Heights School, also commented on the role that their local assessment criteria had played in helping him to get better at using assessment information to guide his teaching. He explained,

1 I think that you get a bit more confidence because of it [the Success Criteria]. You are quite confident telling the parents about what their [child’s] next step is because you know it and it’s clear and you’re working on it. So you are actually doing something about it [their next step], which is good.

(INT: CHS, 2, FST, IND, P. 18, L. 31–33)

Chris felt that his increased confidence in his ability to talk to parents about their child’s next learning steps (lines 2–4) stemmed from his sense of having a degree of control over the meanings that were encapsulated in his school’s Success Criteria. It was because Chris had ownership of these criteria that he felt he could use them to identify (lines 2–3) and act upon (lines 4–5) his students’ next teaching and learning steps. For Chris, involvement in moderation strengthened his confidence in his ability to use student assessment information to identify the implications for future teaching. This provides an indication of the effect that he perceived participating in these processes had upon his assessment competence.

Chapter conclusion

This chapter has shown that those teachers whose experiences of moderation included participating in the development and use of school-specific assessment criteria were the most likely to provide accounts of the ways in which involvement in moderation had strengthened their assessment capability. Drawing on Wenger’s (1998) conception of identity, it has argued that the sense of ownership that teachers garnered from taking part in the co-construction of local assessment criteria increased their confidence in their ability to use these criteria. As well as being the most likely to report that involvement in moderation had strengthened their capacity to arrive at judgements of student achievement, those teachers who took part in the development of local
assessment criteria also described how their experiences of moderation had heightened their ability to use assessment information for formative purposes. The next chapter explores how the findings that have been presented within this thesis contribute to existing understandings about how involvement in moderation affects teachers’ assessment capability.
Chapter 7. Moderation as community forming: Using involvement in moderation to create assessment-focused communities of practice

This study’s results indicate that taking part in social moderation enabled the teachers at each of the participating schools to learn about assessment. Findings also demonstrate that involvement in moderation afforded the teachers at each of these schools qualitatively different learning opportunities. This chapter explores possible reasons for these differences and examines the ways in which this study has contributed to existing understandings about how and what teachers learn about assessment through their involvement in moderation. Using Wenger’s (1998) notion of community, or “learning as belonging” (p. 5), this chapter weaves together the findings from this study and explains why those teachers who used their involvement in moderation to form a community of practice seemed to have the greatest opportunity to learn about assessment from these experiences. To set the scene for this, the chapter begins by reviewing how Wenger’s conception of community was used in this thesis. Next, a summary of the study’s main findings is presented. This is followed by a detailed examination of the ways in which these findings complement existing understandings about how involvement in moderation shapes teachers’ assessment capability. After acknowledging the limitations of this study, the chapter concludes with the presentation of implications for both research and practice.

Characterising a community of practice

As explained in Chapter 2, three dimensions of practice provide a community with its source of coherence (Wenger, 1998). These are mutual engagement, a joint enterprise, and a shared repertoire. This thesis has drawn attention to the ways in which each of these dimensions shaped how and what teachers learned about assessment through their involvement in moderation. For example, it has demonstrated the importance of allocating sufficient time for teachers to mutually engage in moderation activities. Likewise, it has noted that the existence of a clearly defined moderation leadership system seemed to help schools create regular opportunities for teachers to mutually engage in moderation. Similarly, it has identified that the development of a joint enterprise – or a shared perspective about the rationale for participating in moderation – appeared to help teachers sustain their mutual engagement in these activities. Finally, it
has highlighted the ways in which using involvement in moderation to develop local systems and resources (i.e., a shared repertoire) shaped teachers’ learning opportunities. Within this chapter, these three dimensions are brought to the fore and used as a basis for identifying to what degree the teachers at the participating schools were able to use their involvement in moderation to form assessment-focused communities of practice.

**Using involvement in moderation to learn about assessment**

The findings from this study demonstrate that taking part in school-based social moderation activities can provide teachers with valuable assessment-focused professional learning opportunities. This study has shown that teachers used their involvement in moderation to strengthen their ability to collect dependable assessment information. Likewise, evidence was provided that indicated teachers had utilised their experiences of moderation to improve their understandings of the principles and practices of assessment for learning. Although teachers garnered a number of these learning opportunities from using their participation in moderation to discuss judgements of student achievement, most of these opportunities stemmed from the teachers’ involvement in the co-construction of school-specific assessment criteria. To explicate this, the upcoming section opens by examining how this study’s findings augment existing understandings about the learning benefits associated with using involvement in moderation for co-constructing assessment criteria. The following section focuses on those professional learning opportunities that this study has linked with using participation in moderation for talking about judgements of student work. It explores how this study has advanced current understandings about the ways in which involvement in this aspect of moderation can shape teachers’ assessment capability.

*Learning about assessment through involvement in criteria development*

In this study, many of the opportunities that participation in moderation afforded teachers for learning about assessment were linked with their involvement in the co-construction of school-specific assessment criteria. This aspect of the teachers’ moderation practice enabled them to contribute to the development of what Wenger (1998) would describe as a shared repertoire. Where teachers’ experiences of moderation included taking part in the co-construction of local assessment criteria, this shared repertoire included the centrally developed writing standards (Ministry of Education, 2009c, 2010b), their own school’s locally developed assessment criteria, and the shared understandings of these descriptions of achievement that the teachers attained.
The evidence from this study indicates that the experience of participating in the development of school-specific assessment criteria deepened the teachers’ understandings of both the centrally developed writing standards (Ministry of Education, 2009c, 2010b) and their own school’s writing expectations. This helped teachers to get better at selecting appropriate assessment tasks. It also bolstered their confidence in their ability to arrive at dependable judgements of student achievement. As well as enabling teachers to improve their understanding of the role and function of success criteria, taking part in criteria development also provided them with opportunities to get better at crafting effective assessment criteria. Likewise, teachers used their involvement in these processes to broaden their repertoire of strategies for fostering students’ assessment capability. Finally, there was evidence that participating in criteria development helped teachers make connections between judgement making and the identification of next teaching and learning steps. The following section examines what each of these findings contributes to existing understandings about how taking part in moderation shapes teachers’ assessment capability.

This study has shown that teachers were able to use their involvement in the co-construction of school-specific assessment criteria to clarify their understandings of the centrally developed writing standards (Ministry of Education, 2009c, 2010b). Because teachers at both Central Heights and Greenville schools made repeated reference to the centrally developed writing standards as they created their local assessment criteria, they had opportunities to discuss and reach agreement about the meaning of these centrally developed descriptions of achievement. Although Reid (2007) documented the processes that a group of Scottish teachers employed as they collaborated to develop assessment criteria, her study did not provide any insights into the interpretive acts that the teachers engaged in as they used a National Assessment Bank rubric to create more detailed local criteria. In contrast, this thesis provides detailed analyses of the discussions that teachers engaged in as they worked with New Zealand’s centrally developed writing standards (Ministry of Education, 2009c, 2010b) to co-construct local assessment criteria.

The findings reported in this thesis showed how the process of using the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to create their school-specific assessment criteria prompted teachers to draw upon their practitioner knowledge and explore how their personally held expectations related to those that were specified in the centrally developed documents. It demonstrated that when mismatches were identified, these conversations often enabled teachers to align
their personally held standards with those that were stipulated in the centrally developed writing standards. This is a process that Wyatt-Smith et al. (2010) noted did not always occur when teachers in Queensland, Australia were asked to use their centrally developed standards to reach agreement about judgements of student work. They found that some groups of teachers completely disregarded the centrally developed documents and instead relied on unstated personal standards to inform their judgement making. Wyatt-Smith et al. concluded that these teachers assumed that their unstated personal standards were commonly held despite the fact that they “remained unelaborated” (p. 70). In contrast, the findings from the study that is reported on here showed that participating in criteria development prompted teachers to examine and improve the alignment between their personally held standards and those specified in the centrally developed writing standards. This raises the question of whether asking teachers to use centrally developed standards to inform the co-construction of school-specific assessment criteria might increase the likelihood that they engage with and clarify their understandings of these centrally developed definitions of achievement.

Because the teachers in the study that is reported on in this thesis used the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to shape their local assessment criteria, involvement in moderation provided them with opportunities to refine and reach agreement about their understandings of the centrally developed writing standards. These understandings were then encapsulated in each school’s local assessment criteria. Because of this, the school-specific assessment criteria that these teachers developed were closely aligned with New Zealand’s centrally developed writing standards. This is noteworthy because questions have been asked about the extent to which such school-developed resources reflect the content of these centrally developed documents (Ward & Thomas, 2015, 2016). The findings from the current study indicate that teachers took very careful account of the centrally developed writing standards (Ministry of Education, 2009c, 2010b) as they developed their school-specific assessment criteria. This suggests that when such rigorous criteria-development processes are employed, there is good reason to have confidence in the validity of the resulting school-developed resources.

This study also found that engaging in the co-construction of local assessment criteria enabled the teachers at both Central Heights and Greenville schools to develop shared understandings of these descriptions of achievement. This finding is important because teachers need to establish common understandings of standards and assessment criteria to enable them not only to select and develop appropriate assessment tasks (Darr,
2005b) but also to arrive at dependable judgements of student achievement (Klenowski & Wyatt-Smith, 2010a, 2014; Maxwell, 2010). At both Central Heights and Greenville schools, these common understandings provided teachers with the foundation that they required for having conversations about assessment. For example, one group of teachers from Central Heights School used the shared understandings that they had negotiated while developing a set of assessment criteria to identify that there was insufficient alignment between a proposed assessment task and the intended learning domain. This occurred because the experience of developing these criteria allowed the teachers to negotiate a common understanding of the relevant learning domain. As Darr (2005b) indicated, for a validity argument to encompass content considerations, it is necessary to evaluate how well an assessment task represents or samples the intended learning domain. For the teachers at Central Heights School, taking part in the development of local assessment criteria provided them with a sufficiently robust understanding of the relevant learning domain to engage in this process.

This study also demonstrated that the ways in which teachers are grouped for moderation can shape the extent to which their involvement in these processes facilitates the development of shared understandings of standards and assessment criteria. Although it has been noted that New Zealand primary school teachers are grouped in a range of ways to participate in moderation (Hipkins & Robertson, 2012; Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015; Wylie & Hodgen, 2010), the effect that these configurations have upon what teachers learn through their involvement in these activities has received little attention. The findings from this thesis provide evidence that having opportunities to co-construct assessment criteria with members of their teaching teams enabled the teachers at both Central Heights and Greenville schools to develop shared, syndicate-level understandings of these criteria. Yet it was only at Central Heights School, where teachers also had regular opportunities to review, revise, and use their local criteria during full-staff moderation meetings, that teachers developed school-wide understandings of these expectations. At this school, the existence of a clearly defined moderation plan and a commonly understood moderation leadership system played an important role in “enabling engagement” (Wenger, 1998, p. 74) and allowing all teachers to learn about assessment through their involvement in moderation. This finding is important because it extends understandings about how organisational factors, such as ways in which moderation activities are planned, organised, and led, affect the learning opportunities that teachers are afforded.
Participating in the co-construction of school-specific assessment criteria also bolstered teachers’ confidence in their ability to arrive at dependable judgements of student achievement. Building on the work of Reid (2007), who noted that taking part in criteria development seemed to strengthen teachers’ confidence in their judgement making, the results from this thesis showed that teachers explicitly linked their involvement in the development of assessment criteria with having increased confidence in the quality of their judgements of student achievement. This finding also complements the work of Hargreaves et al. (1996), who demonstrated that teachers who had been involved in the development of 7-point rating scales achieved very high levels of inter-rater agreement when they used these scales to arrive at judgements of student work. Considered alongside the work of Reid (2007) and Hargreaves et al. (1996), the findings that have been reported here indicate that providing teachers with opportunities to participate in criteria development could improve the dependability of their judgements of student achievement.

At Central Heights School, involvement in the co-construction of school-specific assessment criteria had the added benefit of enabling teachers to get better at developing these descriptions of achievement. Although some attention has been paid to how taking part in the development of assessment criteria affects the capability that students have to undertake this work (Higgins et al., 1994), the study that is reported on here is the first to examine this in relation to teachers’ involvement in criteria development. It showed that teachers who had the opportunity to work alongside a colleague who possessed a thorough understanding of the function and properties of effective success criteria were able to use their involvement in the co-construction of assessment criteria to strengthen their ability to develop these descriptions of student achievement.

Specifically, as the teachers at Central Heights School proposed, revised, and critiqued their Success Criteria, they learned how important it was to ensure that each criterion reflected an aspect of a centrally developed writing standard. Likewise, they got better at ensuring that each criterion was distinctive and identified a separate facet of the learning domain. Similarly, they came to understand that each criterion needed to describe a measurable or perceivable aspect of the desired performance. Each of these attributes that Brookhart (2013) has identified as “desired aspects of criteria for classroom rubrics” (p. 25). Because involvement in the co-construction of assessment criteria enabled these teachers to think critically about the attributes of a success criterion, their conversations allowed them to deepen their understandings of the
properties of an effective criterion. These skills and knowledge are important because it is generally understood that students need to have a clear understanding of the learning goal that they are working towards in order to achieve it (Absolum et al., 2009; Broadfoot et al., 2002; Sadler, 1989). Whether teachers are sharing teacher-developed success criteria with their students or working alongside them to co-construct such criteria, these descriptions of achievement need to be clearly and effectively written. The findings from the current study indicate that using involvement in school-based moderation processes to co-construct success criteria could provide teachers with opportunities to strengthen their ability to develop well-crafted, effective descriptions of student achievement.

Participating in the co-construction of local assessment criteria also provided the teachers in Greenville School’s middle and senior syndicates with the opportunity to learn a new strategy for fostering students’ assessment capability. These teachers used their involvement in moderation to develop a way of presenting assessment criteria to their students in a form that motivated the children to engage with the criteria. This learning opportunity occurred because one of the participating teachers had extensive knowledge of assessment for learning principles and practices. Because moderation meetings provided a forum for mutually engaging in discussions about assessment ideas, this teacher was able to share his assessment for learning expertise with his colleagues. These teachers arrived at the shared perspective that their child-friendly assessment criteria (which took the form of Must, Should, Could charts) had motivated students to take increased responsibility for their learning. Building on the work of Timperley et al. (2007), who found that teachers require a catalyst or a rationale to motivate them to engage meaningfully in professional learning opportunities, I argued that this commonly held perspective provided these teachers with the rationale that they required to justify their involvement in moderation.

The shared perspective at which the teachers in Greenville School’s middle and senior syndicates arrived also provides an example of what Wenger (1998) has described as the negotiation of a joint enterprise. In this regard, he explained that it is a community’s collectively defined joint enterprise that keeps it together. For the teachers in Greenville School’s middle and senior syndicates, the shared view that their participation in moderation was motivating their students to take increased responsibility for their learning compelled them to try and sustain their involvement in this assessment activity. Accordingly, this was a feature of their participation in moderation that these teachers promoted when they sought to involve their junior-
syndicate colleagues in their moderation approach. Wenger (1998) explained that the enterprise of a community of practice is described as joint, “not in that everybody believes the same thing or agrees with everything, but in that it is communally negotiated” (p. 78). Because Greenville School did not have an agreed moderation leadership system, the teachers in the junior syndicate had not participated alongside their middle- and senior-syndicate colleagues in communally negotiating the Must, Should, Could approach. This might explain why the teachers in the junior syndicate ultimately decided against adopting this system.

These findings reinforce and augment those of Reid (2007), who linked teachers’ involvement in criteria development with the identification of an approach for strengthening students’ assessment capability. In both Reid’s and my studies, the participants used sets of centrally developed writing standards to guide the co-construction of local assessment criteria. Likewise, the teachers in both studies were exposed to information about the principles and practices of assessment for learning. These circumstances appeared to play an important role in enabling the teachers in each study to use involvement in moderation to develop a strategy that promoted students’ assessment capability. Given the lack of attention that this area has received, the findings reported in this thesis lend weight to emerging understandings about how teachers can use their involvement in moderation to support the development of students’ assessment capability. Moreover, findings from the current study also indicate that uncertainty about the leadership of a moderation initiative might have limited the extent to which one group of teachers was able to use their involvement in moderation for learning how to foster students’ assessment capability. Therefore, this thesis also highlights the importance of ensuring that moderation processes are coherently led.

Additionally, this study has provided evidence that those teachers whose experiences of moderation included participating in the co-construction of local assessment criteria were the most likely to use involvement in moderation for formative purposes. One other study of moderation (Hipkins & Robertson, 2012) noted that a group of teachers used their locally developed assessment criteria to identify next learning steps. Hipkins and Robertson did not, however, document the teachers’ involvement in the development of these descriptions of achievement. Likewise, they did not examine how the experience of co-constructing local criteria shaped teachers’ attitudes towards their utility. In contrast, the research reported on here showed that involvement in criteria development played an important role in increasing the likelihood that teachers would use judgement-making conversations for identifying next
teaching and learning steps. Drawing on Wenger’s (1998) conception of identity, I showed that this tendency was linked with the sense of ownership that the teachers developed in relation to their local criteria. Because the teachers had been involved in developing these descriptions of achievement, they felt that they had control over the meanings that they contained. This seemed to increase the teachers’ confidence in their ability to use these criteria for identifying next teaching and learning steps. This suggests that criteria development could be promoted as an activity which increases the likelihood that teachers use involvement in moderation, and indeed judgement making more generally, to inform their teaching practice. This idea is explored in further detail in the next section.

**Learning about assessment through involvement in judgement making**

Assessment information is collected for many purposes (Newton, 2010), but unless teachers use this information for inquiring into and informing their teaching programmes, the assessment process affords students very few direct benefits. The current study found that teachers were able to use their involvement in judgement-making sessions to get better at utilising student assessment information for identifying next teaching and learning steps. It identified that these sessions provided teachers with a forum for examining and critiquing the assessment tasks that they designed and selected. Likewise, it demonstrated that the opportunities that these sessions afforded teachers for mutually engaging in discussions about assessment tasks enabled them to deepen their understandings of factors that can affect the dependability of student assessment information. The following section examines how each of these findings enhances current understandings about the ways in which involvement in moderation contributes to teachers’ assessment capability.

Building on the work of Hipkins and Robertson (2012), who found that teachers used post-moderation discussions to identify next learning steps, the research reported on here showed that involvement in moderation enabled teachers to get better at using judgement-making conversations for formative purposes. This is important because it demonstrates that moderation meetings can provide teachers with opportunities to strengthen this aspect of their assessment capability. Within this study, the teachers at Central Heights School accessed these learning opportunities because they systematically integrated the identification of next teaching and learning steps into their judgement-making processes. This aspect of the teachers’ shared repertoire was founded upon their routine of using their locally developed Success Criteria to inform
conversations about student writing. Because this approach allowed the teachers at Central Heights School to get better at employing judgement-making conversations for identifying next teaching and learning steps, it strengthened their ability to engage in a process that the New Zealand Ministry of Education (2007) has described as learning inquiry.

The evidence from this study indicates that these learning opportunities occurred because one teacher provided her colleagues with the guidance that they required to build the identification of next steps into their judgement-making routine. This teacher, who led her school’s moderation processes, played a very important role in enabling engagement. Factors like this, which make mutual engagement possible, are a crucial feature of any community of practice (Wenger, 1998). With this teacher’s assistance, the identification of next steps became a reflexive aspect of this school’s judgement-making process. It also became the aspect of their involvement in moderation that the teachers at Central Heights School valued the most. In this regard, the shared school-wide perspective that these teachers developed about the benefits of using moderation for inquiring into their teaching coheres with the theory that teachers require a catalyst or a rationale to motivate them to engage meaningfully in professional learning opportunities (Timperley et al., 2007).

At Central Heights School, the shared view that involvement in moderation was helping them inquire into and improve their teaching seemed to enable the teachers to keep their assessment-focused community of practice together. Wenger (1998) stated that a community’s enterprise can never be “fully determined by an outside mandate, by a prescription, or by any individual participant” (p. 80). Although the introduction of the National Standards (Ministry of Education, 2009c) prompted the teachers at Central Heights School to focus their attention on moderation, it was the teachers’ own jointly negotiated response to this situation that provided them with the rationale that they required to sustain their commitment to taking part in moderation. This, in turn, enabled these teachers to continue to learn about assessment through their ongoing involvement in these activities.

These findings are noteworthy because they demonstrate the importance of providing teachers with the guidance and support that they require to communally negotiate a meaningful rationale for engaging in moderation. Although an “individual participant” (Wenger, 1998, p. 80) might not be able to single-handedly prescribe a community’s joint enterprise, the situation at Central Heights School suggests that one member can play an important role in shaping the enterprise that is ultimately
negotiated. Because the teacher who led this school’s moderation processes had a robust understanding of the principles and practices of assessment for learning, she was able to draw her colleagues’ attention to the ways in which they could use their involvement in moderation to strengthen their teaching. Moreover, because this school’s moderation leader scheduled regular opportunities for teachers to take part in full-school moderation meetings, the teachers at Central Heights School were able to use these understandings as the basis for negotiating a school-wide joint enterprise.

This suite of factors enabled the teachers at Central Heights School to communally negotiate the joint enterprise that then allowed them to continue to learn about assessment through their mutual engagement in moderation. This highlights the importance of ensuring that the person who leads a school’s moderation processes has the requisite expertise. The findings from this study indicate that selecting a moderation leader (or, if necessary, employing an external expert) who possesses a thorough understanding of the principles and practices of assessment for learning could increase the likelihood that teachers learn about assessment through their involvement in moderation.

The evidence from this study also demonstrates that taking part in moderation can enable teachers to strengthen their ability to design, select, and administer assessment tasks. Although it has previously been noted that moderation discussions can allow teachers to identify problems with assessment tasks (Hipkins & Robertson, 2012; Klenowski & Wyatt-Smith, 2010b), the research reported on here showed that teachers were able to use their involvement in moderation to deepen their understandings of factors that need to be considered when developing, selecting, and administering assessment activities. For example, involvement in moderation enabled one pair of teachers to learn how differences in assessment conditions can affect student performance. For these teachers, participation in a judgement-making session allowed them to identify how differences in the instructions that they had provided to their students affected the resulting writing. Such differences can pose a threat to validity because the conditions or circumstances of assessment affect the opportunity that students are afforded to demonstrate their skills and abilities (Crooks et al., 1996). In the current study, taking part in a paired judgement-making session enabled two participants to broaden their understanding of those factors that should be considered when designing and administering a common assessment task. This is important because it shows that involvement in moderation can enable teachers to develop the
understandings that they require in order to increase the dependability of the assessment information that they collect.

This study also showed that a group of teachers at Greenville School used their involvement in moderation to examine why students had performed poorly on a locally developed assessment task. The teachers identified that this poor performance was linked to the students’ lack of interest in the assessment activity. Likewise, this study demonstrated that a group of teachers at Central Heights School were able to use their involvement in moderation to diagnose a related issue with a centrally developed assessment task. These teachers felt that this task utilised a context that lacked relevance for many of their students. They feared that the task’s lack of relevance would affect their students’ ability to demonstrate their knowledge and skills in writing.

The concerns that both these groups of teachers raised during their moderation discussions have received considerable attention within the assessment literature. For example, it is understood that both the meaning that children perceive for a task and their interest in it can influence performance (Crooks et al., 1996; Harlen, 1994). Similarly, it has been shown that having limited or negative experiences of the context underpinning a test item can have a detrimental effect on the effort that a student expends on that item (Wise & Smith, 2011). The findings reported in this thesis demonstrate that moderation meetings can provide teachers with a forum for identifying and learning about the complex range of factors that shape the dependability of the assessment information that they gather. These skills are particularly important in contexts like New Zealand, where primary school teachers have considerable freedom to design and select the assessment tasks that they use to collect information about student achievement.

*Using involvement in moderation to support the formation of assessment-focused communities of practice*

This study has shown that the teachers at each of the participating schools were able to use their involvement in moderation to learn about assessment. Yet it has also demonstrated that the teachers at some schools benefited more than others from taking part in these processes. The findings from this study, which are summarised in Table 10, indicate that this variability might be associated with the extent to which involvement in moderation activities enabled teachers to establish an assessment-focused community of practice. Wenger (1998) identified mutual engagement, a joint enterprise, and a shared repertoire as the dimensions of practice that provide a community with its coherence.
The data that are presented in Table 10 indicate that evidence of each of these dimensions was apparent, albeit to varying degrees, in the moderation processes that were observed at two of the participating schools.

Table 10

*Progress by school towards using involvement in moderation to form an assessment-focused community of practice*

<table>
<thead>
<tr>
<th>Dimensions of practice</th>
<th>Evidence of each dimension in teachers’ moderation practice</th>
<th>School</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CHS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SW</td>
</tr>
<tr>
<td>Shared Repertoire</td>
<td>Teachers co-constructed and developed common understandings of local assessment criteria</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Teachers routinely used their local assessment criteria to inform conversations about student work</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Teachers routinely used their conversations about student work to identify next teaching and/or learning steps</td>
<td>•</td>
</tr>
<tr>
<td>Joint Enterprise</td>
<td>Assessment for learning principles provided teachers with a rationale for engaging in moderation</td>
<td>•</td>
</tr>
<tr>
<td>Mutual Engagement</td>
<td>Teachers had frequent, ongoing opportunities to participate in moderation activities</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Teachers had access to a colleague who possessed a thorough understanding of the principles and practices of assessment for learning</td>
<td>•</td>
</tr>
</tbody>
</table>

*Note.* CHS = Central Heights School; GS = Greenville School; RS = Riverside School; FS = Focus syndicate; SW = School wide

These data show that at least two interconnected, assessment-focused communities of practice formed at Central Heights School as a product of their involvement in moderation. The first of these existed at the focus-syndicate level and involved the teachers in the senior syndicate. The second involved all of the teachers...
along with their principal. Although these data indicate that the teachers in Greenville School’s middle syndicate were in the process of establishing a syndicate-level community of practice, their efforts to expand this community to include all the teachers in their school had limited success. Wenger (1998) stated that it is sometimes helpful to think of larger organisations, such as schools, not as single communities of practice but instead as “constellations of interconnected practices” (p. 127). He explained that such organisations are often comprised of different teams that utilise different repertoires and, in some cases, even pursue different enterprises. This helps to account for some of the challenges that the teachers in Greenville School’s middle syndicate encountered as they sought to scale their moderation processes up to include the full school. It also highlights the importance of ensuring that school-based moderation processes are planned and co-ordinated in ways that increase the likelihood that all teachers have the opportunity to develop a shared repertoire and negotiate a joint enterprise.

Although involvement in moderation did not enable the teachers at Riverside School to form an assessment-focused community of practice, this study has shown that some teachers from this school were able to use their experiences of moderation to learn about assessment. This research has, however, demonstrated that involvement in moderation afforded these teachers fewer opportunities to learn about assessment, as compared to their colleagues at either Central Heights or Greenville schools. Appraised collectively, the findings from this study indicate that using involvement in moderation to develop a school-wide, assessment-focused community of practice increases the likelihood that these experiences will afford all teachers opportunities to learn about assessment. Given this, it is argued that the formation of such communities of practice should be a goal of involvement in moderation. After identifying the limitations of this study, a number of recommendations are presented. These recommendations are intended to support the realisation of this goal.

**Limitations of the study**

This study has a number of limitations that might affect the generalisability of its findings. These limitations shape the recommendations that can be made on the basis of this study. They also provide an indication of those areas for which further research on moderation is needed. The first limitation of this study is its small sample size. Because only three schools were involved, any application of findings can only be applied confidently to similar types of schools. An additional limitation of this study is that it was conducted in one moderately sized city in New Zealand’s South Island. Although it
is probable that recommendations from this study’s findings would apply in other areas of New Zealand, and indeed in other countries, this cannot be conclusively ascertained from this research. The duration of this study’s data-collection phase is another limitation of this research. Because this phase only spanned a single school year, it cannot be definitively concluded that involvement in moderation continues to afford teachers opportunities to strengthen their assessment capability.

A further limitation of this study is that no attempt was made to collect data about the level of agreement that involvement in moderation enabled teachers to reach about their judgements of student achievement. Likewise, schools were not asked to provide information that would have shown whether – and if so, to what extent – teachers’ conversations about student writing affected their provisional, pre-moderated judgements of student achievement. Although both of those data sets would have yielded interesting insights, there were a number of reasons that the participating schools were not asked to provide this information. First, a request of this kind would have carried with it assumptions about the nature of the moderation activity in which the teachers were expected to engage. Given that a key goal of this thesis was to examine if and how teachers might access professional learning in assessment through their involvement in the moderation processes that are associated with New Zealand’s National Standards (Ministry of Education, 2009c), it was important to minimise the impact that this study had upon the nature of the learning that the teachers experienced. Second, it was anticipated that such a request might have compromised my ability to build trusting relationships with the participating teachers. As explained in Chapter 2, the introduction of the National Standards signalled a major shift in New Zealand’s assessment and reporting policy (Ministry of Education, 2009d). This shift prompted fears that National Standards assessment data would be used to generate league tables (Croft, 2011). Because of this, many schools felt apprehensive about the collection of such data. This contributed to my decision not to ask the participating schools to provide me with data of that kind.

Another possible limitation of this study is the length of time that has passed since the completion of its data-collection phase. That phase, which began in early 2012, spanned the course of that school year. At that time, many New Zealand primary schools were still adjusting to the introduction of the National Standards (Ward & Thomas, 2012). The schools that participated in this study were no exception and this study’s data-collection phase coincided with each school’s first attempt to engage with moderation within the National Standards context. This meant that many of the learning
opportunities that moderation afforded the participating teachers were associated with their involvement in the development of their school’s moderation approach. Given the time that has elapsed since the data-collection phase, it is expected that most New Zealand primary schools would now have well-established moderation systems. Because of this, the learning opportunities that are described in this study may not be indicative of the opportunities that involvement in moderation currently affords teachers. This does not imply that the insights that this study has provided into how and what teachers learn about assessment through their involvement in moderation are no longer useful. Indeed, the findings associated with the development of moderation resources and approaches could yield particularly useful insights for future practice. This issue is discussed, along with the other implications of this study, in the next two sections.

**Implications for research**

Since data were collected for this study, new assessment resources, such as the Progress and Consistency Tool (PaCT; Ministry of Education, 2014), have been made available to schools. The PaCT has been promoted as a resource that “supports moderation” (Ministry of Education, 2016b, p. 3). This tool includes a series of frameworks that illustrate the learning stages in reading, writing, and mathematics. Teachers have the option of using the PaCT to generate recommendations that they can use to guide their overall judgements of student achievement. Likewise, the government has introduced a new initiative entitled Investing in Educational Success (IES; Parata, 2014). This initiative, which is intended to lift student achievement, has brought about the formation of what have been termed “communities of learning” or “kāhui ako” (Education Review Office, 2016, p. 4). Recent research examining the uptake and early implementation of this initiative has indicated that some schools are using their involvement in these communities to facilitate inter-school moderation processes (Ministry of Education, 2017). Further research is needed to explore how changes in New Zealand’s educational landscape, such as the ones that have just been described, are shaping how and what teachers learn about assessment through their involvement in moderation.

This study has shown that many of the learning benefits that teachers garnered from taking part in moderation were associated with their involvement in the development and use of local assessment criteria. Although the teachers perceived that their involvement in these processes had strengthened their ability to arrive at dependable judgements of student achievement, teachers were not asked to provide a record of their judgements of student work. This meant it was not possible to explore
how taking part in moderation had affected the reliability of their judgements. Further research is needed to examine the relationship between the effects that teachers perceive their involvement in moderation has upon the quality of their judgements and the actual shifts in reliability that participation in moderation leads to. This would provide additional insights into how taking part in these processes shapes teachers’ assessment capability.

This study’s yearlong data-collection phase began shortly after the introduction of the National Standards. Because of this, many of the learning opportunities that moderation afforded the participating teachers were associated with their involvement in the development of moderation systems and resources. Although it is anticipated that these teachers continued to benefit professionally from taking part in moderation, further research is required to definitively ascertain this. Specifically, a longitudinal study is needed to examine whether teachers who are involved in ongoing, established moderation processes continue to access assessment-focused professional learning opportunities from these experiences.

Within the National Standards context, teachers are expected to arrive at and moderate OTJs (Ministry of Education, 2010e). To make or moderate an OTJ for a given student, teachers must appraise a selection of assessment evidence and arrive at an overall judgement about that student’s progress and achievement. With one exception, this study did not provide any insights into how and what teachers learned about assessment through their involvement in the moderation of OTJs. This is because teachers were only observed engaging in conversations about judgements of this kind on one occasion. In contrast, teachers were frequently observed discussing how students had performed on a single, common assessment task. Although recent research examining the implementation of the National Standards has provided some details about how teachers make and moderate OTJs (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015, 2016), that research was questionnaire based rather than observational. Because of that, it allowed few conclusions to be drawn about how teachers’ involvement in the moderation of OTJs shaped their assessment capability.

Given this, more research is needed to investigate how and what teachers learn about assessment through participating in the moderation of OTJs.

**Implications for policy and practice**

This study has generated insights into the contexts, conditions, and activities that enabled New Zealand primary school teachers to use their involvement in the
development of school-based moderation processes to learn about assessment. The recommendations that are presented next draw upon what was found to be successful during this development phase and offer ideas that should enable schools to review and strengthen their current moderation processes. This study found that using involvement in moderation to develop a school-wide, assessment-focused community of practice appeared to increase the likelihood that all teachers had opportunities to learn about assessment. Therefore, the recommendations that are presented next utilise the three dimensions of practice that were shown to enable the teachers at Central Heights School to establish a community of this type. Although the following recommendations provide details that are specific to the New Zealand context, policymakers, school leaders, and teachers in other countries could readily adapt each of these recommendations to suit their individual situations.

**Recommendation 1: Harness the benefits associated with criteria development**

A key recommendation arising out of this study is that the sector should seek to capitalise on the benefits that can be garnered when participation in moderation involves teachers in the co-construction and use of locally developed assessment criteria. For the New Zealand Ministry of Education, this would involve reviewing and clarifying the advice that it has provided to schools about the development and use of such criteria. Currently, the guidance offered to schools about this is ambiguous and potentially contradictory. In late 2014 (A. Carlisle, personal communication, September 11, 2015), the Ministry of Education added a page entitled “Clarifications about National Standards” to its own Assessment Online website (Ministry of Education, n.d.-d). Amidst this information, the statement, “judgements should be made against school-defined standards” was, and indeed still is, included in their list of “common misunderstandings about the National Standards.” It might be inferred from this that the Ministry of Education does not want schools to develop and use local assessment criteria.

Notably, this stance appears to contradict earlier advice that the Ministry of Education provided to schools about the development and use of local criteria. Specifically, one of their own professional learning modules (Ministry of Education, 2011a) specified that teachers’ involvement in moderation could include “using the agreed success criteria informed by the National Standards, the Literacy Learning progressions [sic], Numeracy progressions and stages [sic]” (p. 25) to inform judgements of student achievement. Likewise, at least one other module suggested that
teachers should “collaboratively identify success criteria” (Ministry of Education, 2011b, p. 24). These modules, which were made available shortly after the introduction of the National Standards, are still prominently displayed in the moderation section of the Ministry of Education’s Assessment Online website. Given this, it is unclear how teachers should interpret the more recent advisory that school-defined standards should not be used to inform teachers’ judgements of student achievement (Ministry of Education, n.d.-d).

To address this, the Ministry of Education should update the advice that it has most recently promulgated about the use of school-defined standards. Rather than advising schools not to engage in the development of such local assessment criteria, they should instead seek to provide schools with additional information about the best ways in which to undertake this work. This study has shown that teachers are capable of developing rigorous processes for using the centrally developed writing standards (Ministry of Education, 2009c, 2010b) to co-construct local assessment criteria. Likewise, other research has noted that using locally developed resources to inform moderation processes is commonplace in New Zealand primary schools (Thomas & Ward, 2011; Ward & Thomas, 2012, 2013, 2015, 2016). This suggests that the Ministry of Education could look to schools and teachers to provide them with information about best practice in this area. If, as this thesis has shown, involvement in the development of a resource increases its utility, promoting the co-construction of local assessment criteria may also increase the likelihood that schools make use of the assessment information that these criteria allow them to generate.

Because this study showed that involvement in the development of school-specific assessment criteria provided teachers with such a wide range of opportunities to learn about assessment, a related recommendation is that schools should continue to revise and improve their local criteria. This would enable new teaching staff to establish common understandings of these descriptions of achievement. It would also allow them to develop a sense of ownership of these resources. Given the concerns that have been raised about variations in the local criteria that different schools develop (Ward & Thomas, 2015, 2016), the Ministry of Education’s recently introduced Communities of Learning initiative (Ministry of Education, 2016a) could provide a way forward. Specifically, schools could use their involvement in a community of learning to develop common processes for creating and using local assessment criteria. This would also provide each community with the option of developing a shared pool of assessment criteria. It is anticipated that the sets of criteria within this pool would evolve and
change over time as they were reviewed and adapted by each school in the community. If teachers linked each of these sets of criteria not only with the assessment task or tasks that they had used to assess that learning but also with a selection of examples of students’ responses to these tasks, communities could build up an even richer pool of assessment resources. Additionally, the development of these shared resources would provide the teachers within each community with the requisite framework and understandings to participate meaningfully in inter-school conversations about judgements of student work.

**Recommendation 2: Use assessment for learning principles and practices to link moderation with teaching as inquiry and effective pedagogy**

This study also showed that the teachers who used their involvement in moderation to systematically inquire into their teaching utilised assessment for learning principles and practices to achieve this. On this basis, a further recommendation is that the Ministry of Education should draw more attention to the synergies that exist between moderation, the teaching as inquiry process, and the principles and practices of assessment for learning. This would enable the Ministry of Education to highlight the ways in which moderation processes can be used to support their own conception of “effective pedagogy” (Ministry of Education, 2007, p. 34). Although the relationship between moderation and assessment for learning has been accorded a very brief section on the Assessment Online website (Ministry of Education, n.d.-k), the Ministry of Education does not appear to have established a connection between involvement in moderation and the teaching as inquiry process. Given this, there is considerable scope for providing teachers and schools with more information about the synergies that exist between moderation, teaching as inquiry, and the principles and practices of assessment for learning.

Although some schools in this study were able to integrate assessment for learning principles and practices into their moderation processes, this occurred because at least one teacher at each of these schools possessed the relevant assessment expertise and shared it with colleagues. Given this, efforts should be made to ensure that schools are readily able to gain access to an advisor or facilitator who has assessment for learning expertise. Recent research examining the uptake and early implementation of communities of learning (Ministry of Education, 2017) has identified that many of this initiative’s “across-community teachers” (p. 48) have specialist expertise in assessment. This suggests that communities of learning could present the Ministry of Education with
a way of ensuring that schools have access to the expertise that they require in order to integrate the principles and practices of assessment for learning into their moderation processes. Where a community of learning, or indeed an individual school or kura, does not have internal access to this expertise, it could be encouraged to work with a suitable “accredited facilitator” (Ministry of Education, n.d.-a, para. 1). To assist with this the Ministry of Education should update the information that it currently provides schools in its database of accredited facilitators (Ministry of Education, n.d.-b). Specifically, the terms moderation and/or assessment for learning could be added to the areas of specialisation that are listed within this searchable database.

Recommendation 3: Make time for moderation

The final recommendation that arises from this study relates to the time that schools should reserve for moderation. This study found that those teachers for whom involvement in moderation was both frequent and ongoing had the greatest opportunity to learn about assessment from taking part in these activities. This is congruent with evidence-based research on effective teacher professional learning and development (Timperley et al., 2007). This research identified that those initiatives that were effective usually provided teachers with at least one opportunity to participate per month and typically lasted at least 6 months. It is, therefore, recommended that schools, kura, and (where appropriate) communities of learning, should review their moderation plans and establish an approach that provides all teachers with the opportunity to participate in at least two moderation sessions per term.

These meetings should include at least one criteria-development or review session (henceforth, criteria-development) and one judgement-making session. Given the learning benefits that teachers garnered from having opportunities to participate in both syndicate-level and full-staff moderation meetings, provision should be made for teachers to take part in each of these meeting types. This should support the development of both syndicate-level and school-wide communities of practice. Because this study found that having a clearly defined moderation leadership system seemed to help schools create regular opportunities for teachers to take part in moderation meetings, schools, kura, and communities of learning should ensure that agreement is reached about the teacher or teachers responsible for scheduling and leading their moderation activities.

To enable teachers and students to use their locally developed assessment criteria to guide learning, criteria-development sessions should be scheduled during
non-teaching time or within the first week of a school term. If it is only feasible to schedule one meeting of this type in a given term, then the meeting should involve all teaching staff. If possible, teachers should also use these criteria-development meetings to create or select a common assessment task. This would provide teachers with samples of student work to talk about when they later met to discuss their judgements. These judgement-making sessions should be held during the second half of each term to enable teachers to use the resulting assessment information to identify both what has happened as a result of their instruction and any implications for future teaching.

Closing remarks

A single question provided the catalyst for this study: could providing teachers with opportunities to participate in moderation help the sector fulfil what Absolum et al. (2009) described as an unmet need for teacher professional learning in assessment? This study has shown that the answer to this question is yes. It found that involvement in school-based moderation processes enabled the teachers who participated in this study to strengthen their assessment capability. More specifically, it identified that taking part in moderation provided these teachers with opportunities to build their assessment for learning capability, learn about factors that affect the dependability of assessment information, and develop confidence in their assessment capability. As well as revealing the benefits that teachers garnered from their involvement in moderation, this study also identified some of the challenges that schools encountered as they participated in these processes. These challenges, although not insurmountable, often limited the learning opportunities that participation in moderation afforded teachers. If the sector is interested in meeting the need for teacher professional learning in assessment, and in making the most of the new communities of learning initiative, teachers must be provided with the information, assistance, and time that they require to use involvement in moderation to form assessment-focused communities of practice.
References


https://doi.org/10.1080/01411926.2011.569006


Estyn. (2010). *Evaluation of the arrangements to assure the consistency of teacher assessment in the core subjects at key stage 2 and key stage 3*. Cardiff: Estyn.


Engagement: Elsevier. Retrieved from

https://doi.org/16/j.asw.2007.10.001


understanding of assessment standards and criteria. *Teaching in Higher
Education, 9*(3), 325–335. https://doi.org/10.1080/1356251042000216642

with interview transcription: Towards reflection in qualitative research. *Social

Retrieved from https://education.govt.nz/assets/Documents/Ministry/Investing-
in-Educational-Success/IES-Cabinet-Paper-21-January-2014.pdf


CA: Sage.

development programmes: Final report.* Wellington, New Zealand: Ministry of
Education.


Thrupp, M., & White, M. (2013). *National standards and the damage done: Research, analysis and insight into national standards project* (Report No. 3). Wellington,
New Zealand: New Zealand Educational Institute Te Riu Roa. Retrieved from http://researchcommons.waikato.ac.nz/handle/10289/8394


Dear Chairperson of the Board of Trustees of ____________ School,

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether to participate. If you decide to participate, we thank you. If you decide not to take part, there will be no disadvantage to you of any kind and we thank you for considering our request.

As part of my Doctorate of Philosophy in Education, I am seeking permission for your school to be involved in a study that examines the relationship between the quality of the professional learning opportunities that teachers experience during moderation processes and the trustworthiness of the assessment outcomes that these processes yield. I am seeking permission for your school to be involved in this study as a focus school. This would require involvement from your school during the 2012 school year.

What is the Aim of the Project?
Within the context of the recently introduced National Standards, teachers are required to participate in social moderation processes to improve the trustworthiness or dependability of student assessment information. Although it is widely recognised that participating in such moderation activities can provide teachers with professional learning opportunities, very little is known about the nature and quality of this learning. The proposed study will examine the relationship between the quality of the learning that teachers experience during these moderation processes and the trustworthiness, or dependability, of the assessment information that these moderation processes yield.

What Data or Information will be Collected and What Use will be Made of it?
Although this research project will involve all teachers at your school, it will require the student researcher to develop a particularly close relationship with your school’s senior syndicate. This syndicate will be referred to as the focus syndicate. For most teachers at your school, this project will involve one questionnaire and a series of recorded observations of moderation and meeting sessions. For members of the focus syndicate, however, this research project may also involve an interview or a focus-group session. The student researcher will also collect copies of relevant student assessment information and school-based assessment tools and resources.

Because of the inductive nature of this research, the title and aim of the proposed study, and the research questions that were associated with it, shifted slightly once I was in the field.
All questionnaire results will be coded for later analysis. The principal of your school will also be asked to participate in an interview. Only the student researcher and her Doctoral supervisors will have access to the questionnaire results and any materials collected. All audiotaped moderation sessions and meetings, as well as any audiotaped interviews or focus-group sessions will be transcribed for later analysis. Only the student researcher and her Doctoral supervisors will have access to the audiotaped moderation and meeting sessions, as well as their related transcripts. It is possible that professional transcribers may be employed to transcribe the audiotaped interviews and/or focus-group session. The transcribers will not have access to any audiotaped moderation or meeting sessions. As professionals they guarantee confidentiality and this includes an assurance that any copies of the digital files used for transcription purposes will either be returned or destroyed. The researcher will read, check, and if required edit the accuracy of the transcribed interviews. If it is not possible to employ professional transcribers because of budgetary constraints the researcher will transcribe the interviews and/or focus-group sessions.

Each participant will have the opportunity to read and edit the raw data from the transcribed moderation and meeting sessions that s/he has participated in, as well as any transcripts associated with an interview or focus-group session. This will enable each participant to check that fair and reasonable representation has been made. Pseudonyms will be used to protect the confidentiality of all participants who are referred to.

Although the student researcher will collect information about student achievement, including anonymised copies of student assessment evidence, the focus will be on the teachers’ use of this information and their judgement making processes, and not on the students. Where teachers’ judgements of student achievement or teachers’ use of student assessment evidence are referred to in the unpublished Doctorate or in any published outcomes, it is not expected that this will be in a form that could reasonably be expected to identify the student or students concerned. To protect the identity of your school and its principal, teachers and children, pseudonyms will be used in the development of research findings, in the unpublished Doctorate and in any published outcomes. Copies of all research results will be made available to all participants on request.

The data collected will be securely stored in such a way that only Esther Smaill will be able to gain access to it. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of all participants.

Thank you for reading this information sheet, if you have any further questions regarding this research project please feel free to contact either:

Esther Smaill (Researcher) or Associate Professor Alison Gilmore (Supervisor)
College of Education
03 479 8491 03 479 5036
esther.smaill@otago.ac.nz alison.gilmore@otago.ac.nz

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph. 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
CONSENT FORM FOR BOARDS OF TRUSTEES (FOCUS SCHOOLS)

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:

1. My school’s participation in the project is entirely voluntary;

2. I am free to withdraw my school from the project at any time without any disadvantage;

3. Personal identifying information [student assessment evidence and teacher judgements, questionnaires, audiotaped interviews, moderation, focus-group and meeting sessions, as well as any other materials collected] will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;

4. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of my school and its principal, teachers, and children.

I agree for my school to take part in this project.

.............................................................................  ........................................
(Signature of the Chairperson of the Board of Trustees) (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education  
(Teacher experiences and assessment outcomes)

INFORMATION SHEET FOR PRINCIPALS OF FOCUS SCHOOLS

Dear Principal of ____________ School,

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether to participate. If you decide to participate, we thank you. If you decide not to take part, there will be no disadvantage to you of any kind and we thank you for considering our request.

As part of my Doctorate of Philosophy in Education I am seeking permission for your school to be involved in a study that examines the relationship between the quality of the professional learning opportunities that teachers experience during moderation processes and the trustworthiness of the assessment outcomes that these processes yield. I am seeking permission for your school to be involved in this study as a focus school. This would require involvement from your school during the 2012 school year. Further information about the study and what your school’s involvement will entail are detailed below.

What is the Aim of the Project?
Within the context of the recently introduced National Standards, teachers are required to participate in social moderation processes to improve the trustworthiness or dependability of student assessment information. Although it is widely recognised that participating in such moderation activities can provide teachers with professional learning opportunities, very little is known about the nature and quality of this learning. The proposed study will examine the relationship between the quality of the learning that teachers experience during these moderation processes, and the trustworthiness, or dependability, of the assessment information that these moderation processes yield.

What Data or Information will be Collected and What Use will be Made of it?
Although this research project will involve all teachers at your school, it will require the student researcher to develop a particularly close relationship with either your school’s middle or senior syndicate. If you consent to your school becoming a focus school, further permission would be sought for teachers at your school to be involved in the following ways.

Within each focus school, the student researcher will establish a close relationship with either the middle or senior school syndicate. This syndicate will be referred to as the focus syndicate. The student researcher will establish a relationship with the teacher in this syndicate who has existing responsibility for organising moderation. This teacher will be referred to as the lead teacher.

The Lead Teacher will be asked to:
• Liaise with the student researcher about organisational matters relating to moderation, including the dates and times of scheduled syndicate meetings and moderation sessions, as
well as gaining access to copies of student assessment information and school-based assessment tools and resources. The time commitment associated with this is not expected to exceed one hour per term.

- Complete a questionnaire. It is anticipated that this will take approximately 20 minutes.
- Allow the student researcher to observe and audiotape the moderation sessions that she participates in.
- Allow the student researcher to observe and audiotape her participation in at least two syndicate meetings per term. These syndicate meetings will be in addition to any syndicate meetings that involve a moderation session. For the purposes of this study, a moderation session will be defined as either a meeting or a part of a meeting in which teachers systematically review samples of student work in order to reach agreement about their judgements.
- Allow the student researcher to observe and audiotape her participation in any staff meetings relating to those assessment practices that are relevant to the study.
- Allow the student researcher to make copies of student assessment evidence that is used or referred to during moderation sessions.
- Allow the student researcher to collect information about any judgements of student achievement that are made during moderation sessions.

All teachers in the focus syndicate will be asked to:
- Complete a questionnaire. It is anticipated that this will take approximately 20 minutes.
- Allow the student researcher to observe and audiotape the moderation sessions that s/he participates in.
- Allow the student researcher to observe and audiotape her/his participation in at least two syndicate meetings per term, and other staff meetings as appropriate.
- Allow the student researcher to make copies of student assessment evidence that is used or referred to during moderation sessions.
- Allow the student researcher to collect information about any judgements of student achievement that are made during moderation sessions.

All other teachers at your school will be asked to:
- Complete a questionnaire. It is anticipated that this will take approximately 30 minutes.
- Allow the student researcher to observe and audiotape those moderation sessions that s/he participates in that also involve a member of the focus syndicate.
- Allow the student researcher to observe and audiotape her/his participation in any staff meetings relating to those assessment practices that are relevant to the study.
- Allow the student researcher to make copies of student assessment evidence that is used or referred to during moderation sessions.
- Allow the student researcher to collect information about any judgements of student achievement that are made during moderation sessions.

As the principal of a focus school, you will be asked to:
- Allow the student researcher to observe and audiotape your participation in those moderation sessions that a member of the focus syndicate also participates.
- Allow the student researcher to observe and audiotape your participation in any meetings, including inter-school meetings, relating to those assessment practices that are relevant to the study.

This research project involves: one questionnaire and a series of recorded observations of moderation and meeting sessions. The student researcher will also collect copies of relevant student assessment information and school-based assessment tools and resources. All questionnaire results will be coded for later analysis. Only the student researcher and her Doctoral supervisors will have access to the questionnaire results and any materials collected. The audiotaped moderation and meeting sessions will be transcribed for later analysis. Only the student researcher and her Doctoral supervisors will have access to these audiotaped sessions, as well as their related transcripts. Each participant will have the opportunity to read and edit the
raw data from the transcribed moderation and meeting sessions that s/he has participated in. This will enable each participant to check that fair and reasonable representation has been made. Pseudonyms will be used to protect the confidentiality of all participants who are refereed to.

Although the student researcher will collect information about teachers’ judgements of student achievement and anonymised copies of student assessment evidence, the focus will be on the teachers’ use of this assessment evidence and their judgement making processes, and not on the students. Where teachers’ judgements of student achievement or teachers’ use of student assessment evidence are referred to in the unpublished Doctorate or in any published outcomes, it is not expected that this will be in a form that could reasonably be expected to identify the student or students concerned. To protect the identity of you, your school and its teachers and children, pseudonyms will be used in the development of research findings, in the unpublished Doctorate and in any published outcomes.

In seeking your permission to undertake this research project, I will follow the protocols for collecting information about student achievement, including student assessment evidence, which will be particular to your school site. I am planning to provide information forms to all parents/guardians in your school community, if you also require me to provide additional information forms for children and to obtain signed consent from the children whose assessment evidence may be collected and from their parents/guardians, I am happy to provide forms to enable this to happen. Copies of all research results will be made available to all participants on request.

The data collected will be securely stored in such a way that only Esther Smaill will be able to gain access to it. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of all participants.

Thank you for reading this information sheet, if you have any further questions regarding this research project please feel free to contact either:

Esther Smaill (Researcher) or Associate Professor Alison Gilmore (Supervisor)
College of Education College of Education
03 479 8491 03 479 5036
esther.smaill@otago.ac.nz alison.gilmore@otago.ac.nz

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph. 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education  
(Teacher experiences and assessment outcomes)

CONSENT FORM FOR PRINCIPALS OF FOCUS SCHOOLS

____________ SCHOOL

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I understand that signed consent involves giving the student researcher permission to collect information about student achievement, including copies of student assessment evidence. I am aware that only the student researcher and her Doctoral supervisors will have access to this information.

My school’s protocols for the collection of information about student achievement require, or do not require, the following provisos before the research project can proceed:

1. I require Esther Smaill to provide an information sheet for those children about whom she might collect evidence of achievement. I am aware that this is in addition to the information sheet that she will provide to all the parents/guardians of children at my school.

   Please circle YES OR NO

2. I require Esther Smaill to provide an information sheet for and to obtain a signed consent from those children about whom she might collect evidence of achievement AND to provide an information sheet for and to obtain a signed consent form from their parent/guardians.

   Please circle YES OR NO

I know that:

1. My school’s participation in the project is entirely voluntary;

2. I am free to withdraw my school from the project at any time without any disadvantage;
3. Personal identifying information [student assessment evidence and teacher judgements, questionnaires, audiotaped moderation and meeting sessions, as well as any other materials collected] will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;

4. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of my school, the teachers, the children and myself.

I agree for my school to take part in this project.

.......................................................................................................... ...........................................
(Signature School Principal) (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Dear Principal of ____________ School,

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate, we thank you. If you decide not to take part, there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?
Within the context of the recently introduced National Standards, teachers are required to participate in social moderation processes to improve the trustworthiness or dependability of student assessment information. Although it is widely recognised that participating in such moderation activities can provide teachers with professional learning opportunities, very little is known about the nature and quality of this learning. The proposed study will examine the relationship between the quality of the learning that teachers experience during these moderation processes, and the trustworthiness, or dependability, of the assessment information that these moderation processes yield.

This proposed study seeks to answer the following questions:

1. Within the context of National Standards, what assessment evidence, resources, and processes do teachers use during moderation?
2. How are these materials and processes used to inform and regulate teachers’ judgements of student achievement?
3. What features of effective professional learning can be observed in moderation processes?
4. How do moderation processes affect the dependability of teacher judgements?
5. How do moderation activities affect the dependability of the assessment evidence that teachers use to inform judgement making?
6. What is the relationship between responses to Qs 1-3 and Q4?
7. What is the relationship between responses to Qs 1-3 and Q5?
8. What recommendations emerge from this study that may guide future moderation processes?

This project is being undertaken as part of the requirements for the student researcher’s Doctorate of Philosophy in Education qualification.
What Type of Participants are being sought?
Participation is sought from principals of schools at which teachers are actively participating in moderation processes.

What will Participants be Asked to Do?
Should you agree to take part in this project, you will be asked to:

- Allow the student researcher to observe and audiotape your participation in those moderation sessions that a member of the focus syndicate also participates.
- Allow the student researcher to observe and audiotape your participation in any meetings, including inter-school meetings, relating to those assessment practices that are relevant to the study.

Please be aware that you may decide not to take part in the project without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?
Your participation in this research project will involve a series of observations of moderation and meeting sessions. Only the student researcher and her Doctoral supervisors will have access to the audiotaped moderation and meeting sessions, and the related transcripts. You will have the opportunity to read and edit the raw data from the transcribed moderation and meeting sessions that you have participated in. This will enable you to check that fair and reasonable representation has been made. Pseudonyms will be used to protect the confidentiality of yourself, your colleagues, your school, and any students who are refereed to.

During the preparation of findings for conference presentations and publication the student researcher will take into consideration issues of confidentiality and will use pseudonyms for all participants in the development of research findings and any publications. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of yourself, your colleagues, your school, and any students who are refereed to.

Copies of all research results will be made available on request.

The data collected will be securely stored in such a way that only Esther Smaill will be able to gain access to it. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy. Any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

Can Participants Change their Mind and Withdraw from the Project?
You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What if Participants have any Questions?
If you have any questions about our project, either now or in the future, please feel free to contact either:

Esther Smaill (Researcher) or Associate Professor Alison Gilmore (Supervisor)
College of Education College of Education
03 479 8491 03 479 5036
esther.smaill@otago.ac.nz alison.gilmore@otago.ac.nz

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education  
(Teacher experiences and assessment outcomes)

CONSENT FORM FOR PARTICIPATING FOCUS SCHOOL PRINCIPALS

____________ SCHOOL

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:

1. My participation in the project is entirely voluntary;

2. I am free to withdraw from the project at any time without any disadvantage;

3. Personal identifying information [audiotapes that capture my participation in moderation and meeting sessions] will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;

4. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.

I agree to take part in this project.

.................................................................  ........................................
(Signature of participant)                      (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education
(Teacher experiences and assessment outcomes)

INFORMATION SHEET FOR LEAD TEACHERS

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate, we thank you. If you decide not to take part, there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?
Within the context of the recently introduced National Standards, teachers are required to participate in social moderation processes to improve the trustworthiness or dependability of student assessment information. Although it is widely recognised that participating in such moderation activities can provide teachers with professional learning opportunities, very little is known about the nature and quality of this learning. The proposed study will examine the relationship between the quality of the learning that teachers experience during these moderation processes, and the trustworthiness, or dependability, of the assessment information that these moderation processes yield.

This proposed study seeks to answer the following questions:

1. Within the context of National Standards, what assessment evidence, resources, and processes do teachers use during moderation?
2. How are these materials and processes used to inform and regulate teachers’ judgements of student achievement?
3. What features of effective professional learning can be observed in moderation processes?
4. How do moderation processes affect the dependability of teacher judgements?
5. How do moderation activities affect the dependability of the assessment evidence that teachers use to inform judgement making?
6. What is the relationship between responses to Qs 1-3 and Q4?
7. What is the relationship between responses to Qs 1-3 and Q5?
8. What recommendations emerge from this study that may guide future moderation processes?

This project is being undertaken as part of the requirements for the student researcher’s Doctorate of Philosophy in Education qualification.
What Type of Participants are being sought?
Participation is sought from the teacher or teachers with responsibility for organising school-based moderation activities for the middle-school or senior-school syndicate.

What will Participants be Asked to Do?
Should you agree to take part in this project, you will be asked to:

• Liaise with the student researcher about organisational matters relating to moderation, including the dates and times of scheduled syndicate meetings and moderation sessions, as well as gaining access to copies of student assessment information and school-based assessment tools and resources. The time commitment associated with this is not expected to exceed one hour per term.
• Complete a questionnaire. It is anticipated that this will take approximately 20 minutes.
• Allow the student researcher to observe and audiotape the moderation sessions that you participate in.
• Allow the student researcher to observe and audiotape your participation in at least two syndicate meetings per term. These syndicate meetings will be in addition to any syndicate meetings that involve a moderation session. For the purposes of this study, a moderation session will be defined as either a meeting or a part of a meeting in which teachers systematically review samples of student work in order to reach agreement about their judgements.
• Allow the student researcher to observe and audiotape your participation in any relevant staff meetings.
• Allow the student researcher to make copies of student assessment evidence that is used or referred to during moderation sessions.
• Allow the student researcher to collect information about any judgements of student achievement that are made during moderation sessions.

Please be aware that you may decide not to take part in the project without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?
Your participation in this research project will involve one questionnaire and a series of recorded observations of moderation and meeting sessions. The student researcher will also collect copies of relevant student assessment information and school-based assessment tools and resources. All questionnaire results will be coded for later analysis. Only the student researcher and her Doctoral supervisors will have access to the questionnaire results and any materials collected. All audiotaped moderation and meeting sessions will be transcribed for later analysis. Only the student researcher and her Doctoral supervisors will have access to the audiotaped sessions, as well as their related transcripts. You will have the opportunity to read and edit the raw data from the transcribed moderation and meeting sessions that you have participated in. This will enable you to check that fair and reasonable representation has been made. Pseudonyms will be used to protect the confidentiality of yourself, your colleagues, your school, and any students who are referred to.

During the preparation of findings for conference presentations and publication the student researcher will take into consideration issues of confidentiality and will use pseudonyms for all participants in the development of research findings and any publications. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of yourself, your colleagues, your school, and any students who
are referred to.

Copies of all research results will be made available on request.

The data collected will be securely stored in such a way that only Esther Smaill will be able to gain access to it. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

**Can Participants Change their Mind and Withdraw from the Project?**
You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

**What if Participants have any Questions?**
If you have any questions about our project, either now or in the future, please feel free to contact either:

Esther Smaill (Researcher) or Associate Professor Alison Gilmore (Supervisor)
College of Education College of Education
03 479 8491 03 479 5036
esther.smaill@otago.ac.nz alison.gilmore@otago.ac.nz

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
MODERATION: ASSESSMENT FOR PROFESSIONAL LEARNING IN EDUCATION
(Teacher experiences and assessment outcomes)

CONSENT FORM FOR PARTICIPATING LEAD TEACHERS

____________ SCHOOL

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:

1. My participation in the project is entirely voluntary;

2. I am free to withdraw from the project at any time without any disadvantage;

3. Personal identifying information [student assessment evidence or results, the questionnaire, audiotaped moderation and meeting sessions] will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;

4. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.

I agree to take part in this project.

............................................................................  ........................................
(Signature of participant)  (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education  
(Teacher experiences and assessment outcomes)

INFORMATION SHEET FOR FOCUS SYNDICATE TEACHERS

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate, we thank you. If you decide not to take part, there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?
Within the context of the recently introduced National Standards, teachers are required to participate in social moderation processes to improve the trustworthiness or dependability of student assessment information. Although it is widely recognised that participating in such moderation activities can provide teachers with professional learning opportunities, very little is known about the nature and quality of this learning. The proposed study will examine the relationship between the quality of the learning that teachers experience during these moderation processes, and the trustworthiness, or dependability, of the assessment information that these moderation processes yield.

This proposed study seeks to answer the following questions:

1. Within the context of National Standards, what assessment evidence, resources, and processes do teachers use during moderation?
2. How are these materials and processes used to inform and regulate teachers’ judgements of student achievement?
3. What features of effective professional learning can be observed in moderation processes?
4. How do moderation processes affect the dependability of teacher judgements?
5. How do moderation activities affect the dependability of the assessment evidence that teachers use to inform judgement making?
6. What is the relationship between responses to Qs 1-3 and Q4?
7. What is the relationship between responses to Qs 1-3 and Q5?
8. What recommendations emerge from this study that may guide future moderation processes?

This project is being undertaken as part of the requirements for the student researcher’s Doctorate of Philosophy in Education qualification.
**What Type of Participants are being sought?**
Participation is sought from teachers who are actively involved in their school’s moderation processes.

**What will Participants be Asked to Do?**
Should you agree to take part in this project, you will be asked to:

- Complete a questionnaire. It is anticipated that this will take approximately 20 minutes.
- Allow the student researcher to observe and audiotape the moderation sessions that you participate in.
- Allow the student researcher to observe and audiotape your participation in at least two syndicate meetings per term. These syndicate meetings will be in addition to any syndicate meetings that involve a moderation session. For the purposes of this study, a moderation session will be defined as either a meeting or a part of a meeting in which teachers systematically review samples of student work in order to reach agreement about their judgements.
- Allow the student researcher to observe and audiotape your participation in any relevant staff meetings.
- Allow the student researcher to make copies of student assessment evidence that is used or referred to during moderation sessions.
- Allow the student researcher to collect information about any judgements of student achievement that are made during moderation sessions.

Please be aware that you may decide not to take part in the project without any disadvantage to yourself of any kind.

**What Data or Information will be Collected and What Use will be Made of it?**
Your participation in this research project will involve one questionnaire and a series of recorded observations of moderation and meeting sessions. The student researcher will also collect copies of relevant student assessment information and school-based assessment tools and resources. All questionnaire results will be coded for later analysis. Only the student researcher and her Doctoral supervisors will have access to the questionnaire results and any materials collected. All audiotaped moderation and meeting sessions will be transcribed for later analysis. Only the student researcher and her Doctoral supervisors will have access to the audiotaped sessions, as well as their related transcripts. You will have the opportunity to read and edit the raw data from the transcribed moderation and meeting sessions that you have participated in. This will enable you to check that fair and reasonable representation has been made. Pseudonyms will be used to protect the confidentiality of yourself, your colleagues, your school, and any students who are referred to.

During the preparation of findings for conference presentations and publication the student researcher will take into consideration issues of confidentiality and will use pseudonyms for all participants in the development of research findings and any publications. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of yourself, your colleagues, your school, and any students who are referred to.

Copies of all research results will be made available on request.
The data collected will be securely stored in such a way that only Esther Smaill will be able to gain access to it. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

**Can Participants Change their Mind and Withdraw from the Project?**
You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

**What if Participants have any Questions?**
If you have any questions about our project, either now or in the future, please feel free to contact either:

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>College of Education</th>
<th>Contact Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esther Smaill (Researcher)</td>
<td>Associate Professor</td>
<td>College of Education</td>
<td>03 479 8491 <a href="mailto:esther.smaill@otago.ac.nz">esther.smaill@otago.ac.nz</a></td>
</tr>
<tr>
<td>or Associate Professor</td>
<td>Alison Gilmore (Supervisor)</td>
<td></td>
<td>03 479 5036 <a href="mailto:alison.gilmore@otago.ac.nz">alison.gilmore@otago.ac.nz</a></td>
</tr>
</tbody>
</table>

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education  
(Teacher experiences and assessment outcomes)

CONSENT FORM FOR PARTICIPATING FOCUS SYNDICATE TEACHERS

____________ SCHOOL

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:

1. My participation in the project is entirely voluntary;

2. I am free to withdraw from the project at any time without any disadvantage;

3. Personal identifying information [student assessment evidence or results, the questionnaire, audiotaped moderation and meeting sessions] will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;

4. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.

I agree to take part in this project.

............................................................................  ............................................
(Signature of participant)  (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education
(Teacher experiences and assessment outcomes)

INFORMATION SHEET FOR FOCUS SCHOOL TEACHERS

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate, we thank you. If you decide not to take part, there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?
Within the context of the recently introduced National Standards, teachers are required to participate in social moderation processes to improve the trustworthiness or dependability of student assessment information. Although it is widely recognised that participating in such moderation activities can provide teachers with professional learning opportunities, very little is known about the nature and quality of this learning. The proposed study will examine the relationship between the quality of the learning that teachers experience during these moderation processes, and the trustworthiness, or dependability, of the assessment information that these moderation processes yield.

This proposed study seeks to answer the following questions:

1. Within the context of National Standards, what assessment evidence, resources, and processes do teachers use during moderation?
2. How are these materials and processes used to inform and regulate teachers’ judgements of student achievement?
3. What features of effective professional learning can be observed in moderation processes?
4. How do moderation processes affect the dependability of teacher judgements?
5. How do moderation activities affect the dependability of the assessment evidence that teachers use to inform judgement making?
6. What is the relationship between responses to Qs 1-3 and Q4?
7. What is the relationship between responses to Qs 1-3 and Q5?
8. What recommendations emerge from this study that may guide future moderation processes?

This project is being undertaken as part of the requirements for the student researcher’s Doctorate of Philosophy in Education qualification.
What Type of Participants are being sought?
Participation is sought from teachers who are involved in their school’s moderation processes.

What will Participants be Asked to Do?
Should you agree to take part in this project, you will be asked to:
• Complete a questionnaire. It is anticipated that this will take approximately 20 minutes.
• Allow the student researcher to observe and audiotape those moderation sessions that you participate in that also involve a member of the focus syndicate.
• Allow the student researcher to observe and audiotape your participation in any relevant staff meetings.
• Allow the student researcher to make copies of student assessment evidence that is used or referred to during moderation sessions.
• Allow the student researcher to collect information about any judgements of student achievement that are made during moderation sessions.

Please be aware that you may decide not to take part in the project without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?
Your participation in this research project will involve one questionnaire and a series of recorded observations of moderation and meeting sessions. The student researcher will also collect copies of relevant student assessment information and school-based assessment tools and resources. All questionnaire results will be coded for later analysis. Only the student researcher and her Doctoral supervisors will have access to the questionnaire results and any materials collected. All audiotaped moderation and meeting sessions will be transcribed for later analysis. Only the student researcher and her Doctoral supervisors will have access to the audiotaped sessions, as well as their related transcripts. You will have the opportunity to read and edit the raw data from the transcribed moderation and meeting sessions that you have participated in. This will enable you to check that fair and reasonable representation has been made. Pseudonyms will be used to protect the confidentiality of yourself, your colleagues, your school, and any students who are referred to.

During the preparation of findings for conference presentations and publication the student researcher will take into consideration issues of confidentiality and will use pseudonyms for all participants in the development of research findings and any publications. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve the anonymity of yourself, your colleagues, your school, and any students who are referred to.

Copies of all research results will be made available on request.

The data collected will be securely stored in such a way that only Esther Smaill will be able to gain access to it. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.
Can Participants Change their Mind and Withdraw from the Project?
You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What if Participants have any Questions?
If you have any questions about our project, either now or in the future, please feel free to contact either:

Esther Smaill (Researcher) or Associate Professor Alison Gilmore (Supervisor)
College of Education
03 479 8491 or 03 479 5036
esther.smaill@otago.ac.nz or alison.gilmore@otago.ac.nz

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education
(Teacher experiences and assessment outcomes)

CONSENT FORM FOR PARTICIPATING FOCUS SCHOOL TEACHERS

____________ SCHOOL

I have read the Information Sheet concerning this project and understand what it is
about. All my questions have been answered to my satisfaction. I understand that I am
free to request further information at any stage.

I know that:

1. My participation in the project is entirely voluntary;

2. I am free to withdraw from the project at any time without any disadvantage;

3. Personal identifying information [student assessment evidence or results, the
questionnaire, audiotaped moderation and meeting sessions] will be destroyed at
the conclusion of the project but any raw data on which the results of the project
depend will be retained in secure storage for at least five years;

4. The results of the project may be published and will be available in the University
of Otago Library (Dunedin, New Zealand) but every attempt will be made to
preserve my anonymity.

I agree to take part in this project.

................................................................. .........................................
(Signature of participant) (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any
concerns about the ethical conduct of the research you may contact the Committee through the Human
Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and
investigated and you will be informed of the outcome.
Moderation: Assessment for Professional Learning in Education
(Teacher experiences and assessment outcomes)

INFORMATION SHEET FOR PARENTS/GUARDIANS

My name is Esther Smaill and I am a fully registered primary school teacher. At present, I am studying as a doctoral student at the University of Otago. As part of my study I will be working with the principal and teachers at ____________ School during 2012.

What is the Aim of the Project?
The purpose of my study is to examine what teachers learn professionally during social moderation processes and how this affects the trustworthiness of their assessment decisions. Social moderation is the process of teachers working together with a goal of reaching agreement about their judgements of student work in relation to stated standards. During this project, I will observe and audiotape teachers as they participate in these processes. I will also collect anonymised copies of the student work that they discuss during these sessions. The focus of this study is on the teachers’ judgement making processes and the teachers’ use of assessment evidence and is not in any way focused on the students.

What Data or Information will be Collected and What Use will be Made of it?
I will observe and audiotape teachers as they work together on making judgements about students’ work. Although a discussion of your child’s/children’s achievement may be captured on audiotape, my focus will be on the teachers’ judgement making and the teachers’ use of assessment evidence. Similarly, although I may obtain anonymised copies of your child’s/children’s work, the focus is on how teachers collect and use students’ work to make judgements, and not on the students. To protect the anonymity of your child/children, no student work will be reproduced and no names will be used in the final report.

The audiotapes and transcripts, as well as copies of children’s work will be securely stored in such a way that only I will be able to gain access to them. Five years after the project is finished these materials will be destroyed.

The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand). Copies of all research results will be made available to participants on request.
What if You have any Concerns or Questions?
If you do not want your child’s/children’s work to be involved in the project, there will not be any disadvantage to your child/children of any kind. If, either now or in the future, you have any questions, or would like to request that your child’s/children’s work is not involved in the research project, please contact either:

Esther Smaill (Researcher) or Associate Professor Alison Gilmore (Supervisor)
College of Education College of Education
03 479 8491 03 479 5036
esther.smaill@otago.ac.nz alison.gilmore@otago.ac.nz

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph. 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Appendix B. Success Criteria developed during moderation meetings at Central Heights School

Rubric 1: Personal Experience Writing Success Criteria (Term 1)

<table>
<thead>
<tr>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>After 6 Months</td>
<td>After 1 Year</td>
<td>After 2 Years</td>
<td>After 3 Years</td>
<td>After 4 Years</td>
<td>After 5 Years</td>
<td>After 6 Years</td>
<td></td>
</tr>
<tr>
<td>Can record the initial or dominant sounds in words</td>
<td>Can write several sentences using conjunctions</td>
<td>Spells most high-frequency words correctly</td>
<td>Accurately uses capital letters, full stops, question marks, exclamation marks</td>
<td>Can plan for writing in a variety of ways</td>
<td>Beginning to use complex sentences accurately</td>
<td>Uses complex sentences with interesting beginnings and varied lengths</td>
<td></td>
</tr>
<tr>
<td>Can record some high-frequency words</td>
<td>Can independently spell some high-frequency words correctly</td>
<td>Records dominant sounds in order in words</td>
<td>Accurately records initial consonant blends and endings, approximates medial vowel sounds</td>
<td>Can write for the given purpose independently</td>
<td>Punctuates accurately including speech marks, question marks, commas and apostrophes</td>
<td>Edits to improve writing</td>
<td></td>
</tr>
<tr>
<td>Can read back what they have written</td>
<td>Uses most sentence structures correctly</td>
<td>Beginning to use letters and full stops</td>
<td>Uses upper/lower case letters legibly and correctly</td>
<td>Can write accurate simple and compound sentences</td>
<td>Attempts unknown words</td>
<td>Uses a wide and varied vocabulary, including similes and metaphors</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>After 6 Months</td>
<td>After 1 Year</td>
<td>After 2 Years</td>
<td>After 3 Years</td>
<td>After 4 Years</td>
<td>After 5 Years</td>
<td>After 6 Years</td>
</tr>
<tr>
<td>Can draw in preparation for writing</td>
<td>Can write a sentence that carries a meaningful idea</td>
<td>Records a personal experience, including a personal response</td>
<td>Uses compound sentence structures</td>
<td>Uses vocabulary that clearly conveys ideas and experiences</td>
<td>Uses senses in description in the writing</td>
<td>Uses impact to capture the reader</td>
</tr>
<tr>
<td>Can record personal experiences</td>
<td>Records a simple sentence using a personal voice</td>
<td>Records some compound sentences, including a variety of sentence beginnings and conjunctions</td>
<td>Uses some descriptive language</td>
<td>Adds detail to add interest to writing</td>
<td>Adds detail to the writing by using context-specific language</td>
<td>Uses an introductory and concluding sentence</td>
</tr>
<tr>
<td>Can write a simple sentence</td>
<td>Can think, talk or draw in preparation for writing</td>
<td>Is beginning to edit work to improve meaning</td>
<td>Expresses feelings and ideas</td>
<td>Writes in a personally meaningful way</td>
<td>Involves the reader in the experience by using personal voice</td>
<td>Uses emotive vocabulary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Rubric 2: Recount and Report Writing Success Criteria (Term 2)

<table>
<thead>
<tr>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can record the initial or dominant sounds in words</td>
<td>Can spell essential lists 1 and 2 words accurately</td>
<td>Can write more than 2 sentences using conjunctions</td>
<td>Accurately records initial consonant blends and endings, approximates medial and vowel sounds</td>
<td>Accurately writes a sentence</td>
<td>Can plan for writing in a variety of ways</td>
<td>Beginning to use complex sentences accurately</td>
<td>Uses complex sentences with interesting beginnings and varied lengths</td>
</tr>
<tr>
<td>Can record some high-frequency words</td>
<td>Can read back what they have written</td>
<td>Leaves spaces between words</td>
<td>Beginning to use capital letters and full stops accurately</td>
<td>Can write accurate simple and compound sentences</td>
<td>Can write for the given purpose independently</td>
<td>Punctuates accurately including speech marks, question marks, commas, and apostrophes</td>
<td>Edits to improve writing</td>
</tr>
<tr>
<td>Can read back what they have written</td>
<td>Surface Features</td>
<td>Plans for writing</td>
<td>Identifies words that they are unsure of the spelling, with a circle</td>
<td>Uses upper/lower case letters legibly and correctly</td>
<td>Can write accurate simple and compound sentences</td>
<td>Can write the given purpose independently</td>
<td>Uses a wide and varied vocabulary, including context-specific terminology</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Orally shares an experience</td>
<td>Links 2 or 3 ideas to record an event</td>
<td>Sequences events with a personal comment</td>
<td>Events are recorded in order</td>
<td>Uses the correct tense consistently</td>
<td>Uses a variety of time link words</td>
<td>Uses language and structure appropriate to a report</td>
<td>Uses a range of perspectives</td>
</tr>
<tr>
<td>Retells a simple experience in a sentence</td>
<td>Ideas are sequenced in order</td>
<td>Rereads to check work makes sense</td>
<td>Uses the past tense most of the time</td>
<td>Sequences the events logically</td>
<td>Writes detailed factual, imaginative, and personal recounts</td>
<td>Understands the purpose of a report is to inform</td>
<td>Uses subheadings and paragraphs correctly</td>
</tr>
<tr>
<td>The story matches the picture</td>
<td></td>
<td></td>
<td>Gives an introduction</td>
<td>Gives an interesting and informative introduction</td>
<td>Uses a relevant dynamic introduction and concluding statements</td>
<td></td>
<td>Writes creative recounts from a range of perspectives</td>
</tr>
</tbody>
</table>
Rubric 3: Character Writing Success Criteria (Term 3)

<table>
<thead>
<tr>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
<th>Surface Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>After 6 Months</td>
<td>After 1 Year</td>
<td>After 2 Years</td>
<td>After 3 Years</td>
<td>End Year 4</td>
<td>End Year 5</td>
<td>End Year 6</td>
<td></td>
</tr>
<tr>
<td>Can record the initial or dominant sounds in words</td>
<td>Plans for writing</td>
<td>Plans for writing</td>
<td>Plans for writing</td>
<td>Can plan for writing in a variety of ways</td>
<td>Can plan for writing in a variety of ways</td>
<td>Can plan for writing in a variety of ways</td>
<td></td>
</tr>
<tr>
<td>Can record some high-frequency words</td>
<td>Can spell essential lists 1 and 2 words accurately</td>
<td>Identifies words that they are unsure of the spelling, with a circle</td>
<td>Accurately use capital letters, full stops, question marks, and exclamation marks</td>
<td>Can write for the given purpose independently</td>
<td>Can write for the given purpose independently</td>
<td>Can write for the given purpose independently</td>
<td></td>
</tr>
<tr>
<td>Can read back what they have written</td>
<td>Can write more than 2 sentences using conjunctions</td>
<td>Accurately records initial consonant blends and endings, approximates medial and vowel sounds</td>
<td>Uses upper/lower case letters legibly and correctly</td>
<td>Can write accurate simple and compound sentences</td>
<td>Can write accurate simple and compound sentences</td>
<td>Can write accurate simple and compound sentences</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leaves spaces between words</td>
<td>Beginning to use initial blends</td>
<td>Use the character’s emotions and the character</td>
<td>Uses upper/lower case letters legibly and correctly</td>
<td>Uses upper/lower case letters legibly and correctly</td>
<td>Uses upper/lower case letters legibly and correctly</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Beginning to record initial blends</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Deep Features

<table>
<thead>
<tr>
<th>Deep Features</th>
<th>Deep Features</th>
<th>Deep Features</th>
<th>Deep Features</th>
<th>Deep Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>As a character</td>
<td>Uses descriptive language to describe the character’s physical characteristics</td>
<td>Uses descriptive language to describe the character’s personality</td>
<td>Uses descriptive language to describe the character’s emotions and the character</td>
<td>Uses descriptive language to describe the character’s behaviours specifically</td>
</tr>
<tr>
<td>Can describe what the character looks like</td>
<td>Uses descriptive language to describe the physical characteristics</td>
<td>Includes a personal comment</td>
<td>Uses a variety of interesting descriptive language</td>
<td>Uses more complex sentences to frame the ideas</td>
</tr>
<tr>
<td>Can describe what the character does</td>
<td>Uses descriptive vocabulary to describe a character’s personality</td>
<td>Can describe a range of features and behaviours of the character</td>
<td>Includes a personal comment</td>
<td>Shows personal voice and makes the character come alive</td>
</tr>
<tr>
<td>The picture matches the character or story told about the character</td>
<td>Is beginning to describe the character</td>
<td>Uses different sentence starters not just the character’s names or pronouns over again</td>
<td>Uses different sentence starters not just the character’s names or pronouns over again</td>
<td>Shows the character by giving detailed examples of behaviour and personality</td>
</tr>
</tbody>
</table>

Deep Features

<table>
<thead>
<tr>
<th>Deep Features</th>
<th>Deep Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses a wide vocabulary including adjectives and similes</td>
<td>Uses a wide vocabulary including adjectives and similes</td>
</tr>
<tr>
<td>Shows personal voice and makes the character come alive</td>
<td>Shows personal voice and makes the character come alive</td>
</tr>
<tr>
<td>Can organise ideas into paragraphs about different aspects of the character</td>
<td>Can organise ideas into paragraphs about different aspects of the character</td>
</tr>
</tbody>
</table>

Deep Features

<table>
<thead>
<tr>
<th>Deep Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses complex sentences with interesting beginnings and varied lengths</td>
</tr>
<tr>
<td>Edits to improve writing</td>
</tr>
<tr>
<td>Uses a wide and varied vocabulary, including context-specific terminology</td>
</tr>
</tbody>
</table>
## Rubric 4: Narrative Writing Success Criteria (Term 4)

<table>
<thead>
<tr>
<th>After 6 Months</th>
<th>After 1 Year</th>
<th>After 2 Years</th>
<th>After 3 Years</th>
<th>End Year 4</th>
<th>End Year 5</th>
<th>End Year 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surface Features</strong></td>
<td><strong>Surface Features</strong></td>
<td><strong>Surface Features</strong></td>
<td><strong>Surface Features</strong></td>
<td><strong>Surface Features</strong></td>
<td><strong>Surface Features</strong></td>
<td><strong>Surface Features</strong></td>
</tr>
<tr>
<td>Can record the initial or dominant sounds in words</td>
<td>Plans for writing</td>
<td>Plans for writing cooperatively</td>
<td>Accurately use capital letters, full stops, question marks, and exclamation marks</td>
<td>Can plan for writing in a variety of ways</td>
<td>Beginning to use complex sentences accurately</td>
<td>Uses complex sentences with interesting beginnings and varied lengths</td>
</tr>
<tr>
<td>Can record some high-frequency words</td>
<td>Can spell essential lists 1 and 2 words accurately</td>
<td>Identifies words that they are unsure of the spelling, with a circle</td>
<td>Uses upper/lower case letters legibly and correctly</td>
<td>Can write for the given purpose independently</td>
<td>Punctuates accurately including speech marks, question marks, commas, and apostrophes</td>
<td>Edits to improve writing</td>
</tr>
<tr>
<td>Can read back what they have written</td>
<td>Can write more than 2 sentences using conjunctions</td>
<td>Accurately records initial consonant blends and endings, approximates medial and vowel sounds</td>
<td>Correctly spells high-frequency words: essential lists 1–4</td>
<td>Can write accurate simple and compound sentences</td>
<td>Attempts unknown words and uses a dictionary to check accuracy</td>
<td>Uses a wide and varied vocabulary, including context-specific terminology</td>
</tr>
<tr>
<td>Can use a capital letter and a full stop</td>
<td>Leaves spaces between words</td>
<td>Beginning to use capital letters and full stops accurately</td>
<td>Edits own work for sense and accuracy</td>
<td>Uses correct tense consistently</td>
<td>Deep Features (End Year 6)</td>
<td></td>
</tr>
<tr>
<td>Can leave spaces between words</td>
<td><strong>Deep Features</strong></td>
<td><strong>Deep Features</strong></td>
<td><strong>Deep Features</strong></td>
<td><strong>Deep Features</strong></td>
<td><strong>Deep Features</strong></td>
<td><strong>Deep Features</strong></td>
</tr>
<tr>
<td><strong>Deep Features</strong></td>
<td>Can relate to own event</td>
<td>Retells a known story using personal voice</td>
<td>Follows their story plan</td>
<td>Logically sequences ideas</td>
<td>Ideas are sequenced in paragraphs</td>
<td></td>
</tr>
<tr>
<td>Can write a simple sentence story</td>
<td>Writing matches a picture</td>
<td>Uses descriptive language to tell the story</td>
<td>Introduces a problem and resolution in a story</td>
<td>Uses a wider descriptive vocabulary including simile</td>
<td>Can write a range of genre in narrative</td>
<td></td>
</tr>
<tr>
<td>Can relate to own event</td>
<td>Beginnings of a story structure</td>
<td>Writes relevant sentences and adds detail</td>
<td>Uses a wider vocabulary</td>
<td>Writes for different perspectives 1st person, 3rd person</td>
<td>Writes more complex storylines with more characters</td>
<td></td>
</tr>
<tr>
<td>Writing matches a picture</td>
<td>Sequences ideas into a beginning, middle and end</td>
<td>Begins to use direct speech</td>
<td>Begins to carry the storyline with direct speech</td>
<td>Uses specific adjectives and adverbs to describe character and setting</td>
<td>Vivid vocabulary used</td>
<td></td>
</tr>
<tr>
<td>Can use imagination</td>
<td>Can use imagination</td>
<td>Includes setting and character detail</td>
<td>Engages the reader with title and impact</td>
<td>Uses a range of sentence lengths</td>
<td>Engages the reader with title and impact</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix C. Sample of a Must, Should, Could chart

### Writing End Year 4

**Must**
- Correctly use...
  - Capital letters
  - Full stops
  - Question marks
  - Exclamation marks
  - Text structure
- Carefully choose appropriate vocabulary
- Use a variety of sentence beginnings
- Use some compound sentences
- Use similes
- Use a dictionary
- Use and spell words with common prefixes and suffixes correctly
- Re-read and improve text during draft writing
- Proofread for accuracy of spelling and punctuation

**Should**
- Use...
  - Speech marks “ ”
  - Commas ,
  - Apostrophes for contractions
  - Some complex sentences
  - A thesaurus

**Could**
- Use...
  - Commas to denote clauses
  - Possessive apostrophes

---

TEACHER: Subject, verb agreement / Pronouns, prepositions / Tense agreement / Vary structure and length of sentences / Onomatopoeia / Alliteration / Correctly spell Essential Lists 1-4 and some of lists 5-7 / Spelling rules and conventions / Grammar
Appendix D. Excerpt from field notes

Meeting 1: Greenville School
Meeting Type: Middle & senior syndicate meeting
Date: 3 April, 2012, 3:15 pm
Location: Middle syndicate meeting room
Present: Susan, Sophia, Phillip, Jack, Richard, Jessica, Anna

Seating positions

Pre-meeting conversation

- Jack, Richard, Jessica = Senior syndicate
- Anna = Special Education Needs Co-ordinator (SENCO)
- Phillip, Sophia (teaching 5 years), Susan = Middle syndicate
- Phillip has been in NZ for 4 years. He is from England. Commented that NZ is like a breath of fresh air after the UK (Key stage testing etc.)
- Middle and senior syndicate previously worked together. They have split into two syndicates more recently because the number of classes/teachers grew.

Formal meeting session

Susan: Questions whether to develop criteria for specific genre or for what a child’s work should look like in general terms.

Phillip: Talks about the issue of a child getting a good checklist “score” for a piece of writing that is actually poor quality.

*Think about how Phillip’s comment relates to Sadler’s writing. Check his article: Indeterminacy in the use of preset criteria for assessment and grading.*

3:30

Jack: Questions how much scaffolding to provide students during an assessment. Doesn’t want to have to “step out” of his teaching mode too much and give over too much time to assessing.

Susan: Acknowledges the issue of teacher assistance.

Anna: States that an assessment would be unsupported writing.
### Appendix E. Overview of meetings involving teachers in the focus syndicate at Central Heights School

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Date</th>
<th>Meeting Type</th>
<th>Facilitator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14-02-12</td>
<td>Focus syndicate</td>
<td>Kerry</td>
<td>Drafting Success Criteria for personal experience writing.</td>
</tr>
<tr>
<td></td>
<td>Term 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15-02-12</td>
<td>Full staff (including Principal)</td>
<td>Kerry</td>
<td>Reviewing and revising Success Criteria for personal experience writing.</td>
</tr>
<tr>
<td></td>
<td>Term 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>28-02-12</td>
<td>Focus syndicate</td>
<td>Kerry</td>
<td>Discussing judgements of students’ personal experience writing.</td>
</tr>
<tr>
<td></td>
<td>Term 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>04-04-12</td>
<td>Full staff (including Principal)</td>
<td>Kerry</td>
<td>Discussing judgements of students’ personal experience writing.</td>
</tr>
<tr>
<td></td>
<td>Term 1</td>
<td>*William on leave</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td>Focus syndicate</td>
<td></td>
<td>Drafting Success Criteria for recount and report writing.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*Not observed, but reference is made to this meeting in a subsequent transcript</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>09-05-12</td>
<td>Full staff (including Principal)</td>
<td>Kerry</td>
<td>Reviewing and revising Success Criteria for recount and report writing.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>19-06-12</td>
<td>Focus syndicate</td>
<td>Kerry</td>
<td>Discussing judgements of students’ recount and report writing.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>27-06-12</td>
<td>Full staff</td>
<td>Kerry</td>
<td>Discussing judgements of students’ recount and report writing.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>25-07-12</td>
<td>Full staff (including Principal)</td>
<td>Kerry</td>
<td>Full-school judgement-making session using an e-asTTle student exemplar and the e-asTTle rubrics.</td>
</tr>
<tr>
<td></td>
<td>Term 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meeting</td>
<td>Date</td>
<td>Meeting Type</td>
<td>Facilitator</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>9(^13)</td>
<td>30-08-12</td>
<td>Focus syndicate: Paired meeting involving William &amp; Margaret</td>
<td>William &amp; Margaret</td>
<td>Discussing judgements of students’ character writing.</td>
</tr>
<tr>
<td>10</td>
<td>07-09-12</td>
<td>Focus syndicate: Paired meeting involving Chris &amp; Kerry</td>
<td>Kerry</td>
<td>Discussing judgements of students’ character writing.</td>
</tr>
<tr>
<td>11</td>
<td>19-09-12</td>
<td>Full staff (including Principal)</td>
<td>Kerry</td>
<td>Discussing judgements of students’ character writing.</td>
</tr>
<tr>
<td>12</td>
<td>24-10-12</td>
<td>Full staff (including Principal)</td>
<td>Kerry</td>
<td>Devising Success Criteria for narrative writing.</td>
</tr>
<tr>
<td></td>
<td>Term 4</td>
<td>Focus syndicate</td>
<td></td>
<td>Reviewing Success Criteria for narrative writing.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Not observed, but this meeting was scheduled during Meeting 12</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Term 4</td>
<td>Full staff</td>
<td></td>
<td>Discussing judgements of students’ narrative writing.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>I was invited to this meeting but was unable to attend it</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^13\) Meetings 9 and 10 are numbered separately but as each member of the focus syndicate was only involved in one of these paired sessions, they are counted elsewhere as a single meeting.
## Appendix F. Overview of meetings involving teachers in the focus syndicate at Greenville School

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Date</th>
<th>Meeting Type</th>
<th>Facilitator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>03-04-12</td>
<td>Inter-syndicate</td>
<td>Susan</td>
<td>Discussing how to approach moderation and the resources that they will need to develop for this purpose. The “Must, Should, Could” concept is introduced.</td>
</tr>
<tr>
<td></td>
<td>Term 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>08-05-12</td>
<td>Inter-syndicate</td>
<td>Susan</td>
<td>Discussing the development of school-specific moderation resources. Susan shares two drafts of charts that she has developed: a Must, Should, Could chart and a chart detailing the characteristics of a recount. It is decided that the charts for each genre/writing purpose will be presented as WALTs/WYLFs. A decision is made to devise Must, Should, Could charts for other year levels.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15-05-12</td>
<td>Focus syndicate</td>
<td>Susan</td>
<td>Devising Must, Should, Could charts for students in years 2 &amp; 4.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>30-05-12</td>
<td>Inter-syndicate</td>
<td>Susan</td>
<td>Discussing judgements of students’ recount writing.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>06-06-12</td>
<td>Full staff (including Principal)</td>
<td>Susan</td>
<td>Introducing the moderation system that the middle and senior syndicates have developed to junior syndicate colleagues. Seeking buy-in from these colleagues.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>13-06-12</td>
<td>Inter-syndicate</td>
<td>Susan</td>
<td>Developing WALT/WYLF charts for selected writing purposes/genres.</td>
</tr>
<tr>
<td></td>
<td>Term 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Term 3</td>
<td>Inter-school</td>
<td></td>
<td>Inter-school judgement-making session.</td>
</tr>
</tbody>
</table>

I was invited to this meeting but was unable to attend it. Susan was the only member of the focus syndicate to attend this meeting.
### Appendix G. Overview of meetings involving teachers in the focus syndicate at Riverside School

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Date</th>
<th>Meeting Type</th>
<th>Facilitator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20-04-12</td>
<td>Focus syndicate</td>
<td>Stephen</td>
<td>Discussing judgements of students’ recount writing.</td>
</tr>
<tr>
<td></td>
<td>Term 1</td>
<td>Not observed, but reference was made to this meeting during Meeting 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>14-08-12</td>
<td>Full staff (including Principal)</td>
<td>Principal</td>
<td>Discussing judgements of students’ recount writing, using both the asTTle recount rubric and the National Standards.</td>
</tr>
<tr>
<td></td>
<td>Term 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>27-09-12</td>
<td>Focus syndicate</td>
<td>Stephen</td>
<td>Reviewing assessment criteria and task administration in preparation for school-wide assessment of functional/instruction writing using three NEMP tasks.</td>
</tr>
<tr>
<td></td>
<td>Term 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>23-10-12</td>
<td>Full staff (including Principal)</td>
<td>Principal</td>
<td>Discussing issues that arose from using NEMP assessment tasks.</td>
</tr>
<tr>
<td></td>
<td>Term 4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix H. Interview guide for principals

Greeting and thank you

Purpose of moderation

The term “moderation” can mean different things to different people….

1. What does the term moderation mean to you?

Prompt: What steps and processes does moderation involve?

2. What, in your opinion, are the primary purposes of moderation?

3. How well do you think that these purposes are currently being met at your school?

4. Do you have any goals or next steps planned in terms of your school’s moderation processes?

5. I’m interested in hearing about how you see your role in relation to your school’s moderation processes.

Benefits and challenges

6. What professional benefits, if any, have you gained personally through your involvement in _______ School’s moderation processes?

7. How, if at all, do you think that moderation is benefiting your school?

8. What, if anything, have you personally found challenging or hard about participating in moderation?

9. What do you think are the main challenges that moderation poses for schools?

Practices prior to 2010

10. Prior to the introduction of National Standards, did teachers at your school meet to moderate student work in relation to school benchmarks? If yes:

a. Can you tell me about this?

b. Which curriculum or learning areas was this done for?

c. Had your school received any external support, or professional development, to help establish these “moderation” systems? If yes:

i. Can you describe the support that your school received?

d. How, if at all, have your school’s moderation processes changed since the introduction of National Standards?
Practices since 2010

11. Since the introduction of National Standards, has your school received any support from external facilitators or advisors to assist with moderation processes?

If YES:

    a. Can you describe the assistance that you received?
    b. How adequate was this support?

If NO:

    c. How did your school establish the moderation processes that it is using currently?

Professional development: Assessment

12. Within the last 5 years, has your school participated in any professional development contracts that have included a focus on assessment?

Prompt: For example, Assess to Learn (AToL), The Literacy Professional Development Project (LPDP).

13. How, if at all, do you think that participating in this contract informed your school’s moderation processes?

14. Is there anything further that you would like to add?

THANK YOU VERY MUCH FOR TAKING THE TIME TO TALK WITH ME.
Appendix I. Interview guide for focus syndicate and lead teachers

Greetings and thank you

Section 1: Meaning and purpose of moderation

1. What does the term moderation mean to you?
   Prompt: What steps and processes does moderation involve?

2. What, in your opinion, are the primary purposes of moderation?

3. You and your colleagues have worked really hard this year to establish your moderation processes. How well do you think that the moderation purposes that you have identified are being met at your school?

Section 2: Meeting Types

4. So far this year, you have participated in a variety of different moderation meetings. You have contributed to (read school-specific meeting-type cards and place cards on the table). Are there any meeting types that I’ve missed? If yes: Record meeting type on a card and ask the teacher to tell me a little about it.

I’m really interested in hearing about your judgement-making processes, and I was wondering if you could talk to me about how your involvement in all these moderation sessions (gesture to cards) has informed your judgement making.

5. Have some of these meeting types (gesture to cards) been more useful than others in terms of informing or shaping your judgement-making processes? If yes:
   a) Can you please tell me more about this?

6. Still thinking about all the (gesture to cards) moderation processes that you have participated in this year, how have these sessions informed your assessment practice?

Prompts:

• How has moderation affected the ways that you plan for or conduct assessments?
  OR

• How has moderation made you reflect on or adjust the way that you plan for or conduct assessments?
7. Have some of these meeting types been more useful than others in terms of informing or shaping your **assessment practice**? If yes:
   a) Can you please tell me more about this?

8. Still thinking about all the (gesture to cards) moderation processes that you have participated in this year, how have these sessions informed your **teaching practice**?

9. Have some of these meeting types been more useful than others in terms of informing or shaping your **teaching practice**? If yes:
   a) Can you please tell me more about this?

**Section 3: Judgement Making**

10. Thinking specifically about the moderation meeting that you took part in on ________, did any of your judgements change or shift because of/during this session?
   a) Can you talk about how and why your judgement(s) changed?
   b) How might this affect the way that you make judgements from now on?

11. Reflecting on other **moderation sessions** in which you discussed your judgements of student work, did participating in this type of meeting prompt you to adjust any judgements of student achievement?
   a) Can you explain how and why your judgement(s) changed?
   b) How did this affect the way that you made judgements from then on?

12. We’ve talked about the processes that your school has put in place formally to enable teachers to moderate their judgements of student work in writing. Have you been involved in or had any other experiences that have informed your judgement making?

**Section 4: Resources**

13. Developing and clarifying understandings of descriptions of student achievement is an important part of the moderation process.

There are a variety of resources available that provide **descriptions of student achievement in writing**; these are just some of them (place the National Standards, Literacy Learning Progressions, New Zealand Curriculum and e-asTTle rubric on the table).

I’m interested in hearing about how participating in moderation has informed or clarified your understandings of such descriptions of student achievement in writing.
14. Which resources have you found **most useful** during moderation sessions?
   a) Can you please tell me more about this?

15. Which resources have you found **least useful** during moderation sessions?
   a) Can you please tell me more about this?

16. You have taken part in a variety of different moderation activities this year (place moderation-type cards on the table again).

   Have some of these meeting types been more useful than others in terms of informing your understanding of the **descriptions** of writing behaviour that are provided in these, and other, documents? If yes:
   a) Can you please tell me more about this?

**Section 5: Reporting**

17. How has participating in moderation processes informed the ways in which you report to parents?

18. How has participating in moderation processes informed the ways in which you report to the Board of Trustees?

**Section 6: Benefits, challenges, and ideas**

19. Thinking about **all** the moderation processes you have participated in this year, what aspects of moderation have you found **hard or challenging**?

20. Have some of these meeting types been more **difficult or challenging** than others? Can you please tell me more about this?

21. Thinking about **all** the moderation processes you have participated in this year (gesture to cards), what aspects of moderation have you found **professionally useful or rewarding**?

22. Have some of these meeting types been more **useful or professionally rewarding** than others? Can you please tell me more about this?

23. Have you got any ideas or suggestions as to how moderation could be made more professionally rewarding at _________ School?

24. Is there anything further that you would like to add about your experiences of moderation?

**THANK YOU VERY MUCH FOR TAKING THE TIME TO TALK WITH ME.**
Appendix J. Interview guide for additional lead teacher questions

1. Can you tell me about how and why you took responsibility for moderation at your school?
2. What are your responsibilities as the lead teacher for moderation at your school?
3. Do you receive any release time or management units to assist or support the work you do in this role? If yes:
   a. Can you describe the support that you receive?
   b. How adequate is this support?
4. Thinking about all the moderation processes (gesture to cards) you have taken part in and facilitated this year, what aspects of these moderation meetings have you found hard or challenging?
   a. Have some of these meeting types been more difficult or challenging to facilitate than others? If yes:
      i. Can you please tell me more about this?
   b. Have some of these meeting types been more difficult or challenging to participate in than others? If yes:
      i. Can you please tell me more about this?
5. How did your school establish the moderation processes that it is using at the moment? (Evolution from prior/existing systems?)
6. Are any aspects of your school’s moderation processes recorded? If yes:
   a. What records are kept?
   b. How are these records used?

THANK YOU VERY MUCH FOR TAKING THE TIME TO TALK WITH ME.
Appendix K. Teacher questionnaire

Reference Number 11/296
16 December 2011

**Moderation: Assessment for Professional Learning in Education**
(teacher experiences and assessment outcomes)

**TEACHER QUESTIONNAIRE**

1. Name: ____________________________
2. School: ___________________________
3. How many years have you taught at this school? ____________________________
4. What year level(s) do you currently teach? ____________________________
5. How many years have you been teaching this age group? ____________________________
6. How many years teaching experience do you have in total? ____________________________

**Practice before 2010**

7. Prior to the introduction of National Standards, did you meet with colleagues to moderate student work/assessments in relation to school benchmarks?
   - Yes
   - No

If yes, please answer Qs 8–13. If no, proceed to Q14.

8. Please indicate where these sessions took place. Please tick one box only.
   a. At the school you currently work. □
   b. At a school where you previously worked. □
   c. At both your current and previous schools. □

9. Which curriculum or learning areas did these sessions take place for?

   __________________________________________
   __________________________________________
   __________________________________________

10. How often did meet for this purpose? __________________________________________
    __________________________________________
11. Briefly describe what happened at these meetings.  


12. Prior to the introduction of National Standards, did your school work with an external facilitator or advisor to develop the approach that was used when teachers met to moderate student work/assessments in relation to school benchmarks?  
Yes  No  
If yes:  

13. Please explain if and how this was helpful.  


Practice since 2010  

14. Since the introduction of National Standards, has an external facilitator or advisor worked with your school to support the development of your moderation processes?  
Yes  No  
If yes, please answer Qs 15-17. If no, proceed to Q18.  

15. Please indicate what kind of assistance you received. Please tick all the relevant boxes.  
   a. General assistance with moderation.  
   b. Assistance with moderation processes for writing.  
   c. Assistance with moderation processes for reading.  
   d. Assistance with moderation processes for mathematics.  
16. Please describe the assistance that your school received.  


17. Please explain if and how this assistance was helpful.  


18. Since the introduction of National Standards, have you attended any courses or workshops that have informed your moderation practice?  
Yes  No  
If yes, please answer Qs 19-20. If no, proceed to Q21.
19. Please describe the course(s) or workshop(s) you attended.

__________________________________________________________

20. Please explain if and how helpful you found each course.

__________________________________________________________

Current practice

21. In your opinion, what are the main purposes of the moderation activities that you currently participate in?

__________________________________________________________

22. What is your principal’s role in your school’s current moderation processes?

__________________________________________________________

Please refer to the scale below and indicate to what extent you agree with the following statements.


23. Moderation is for teachers to reach agreement about the criteria that will be used to assess student work.

24. Moderation is for teachers to reach consensus about their judgements of student achievement.

25. Moderation is for teachers to share, review and refine their judgement making processes.

26. Moderation is for teachers to plan, review and refine assessment tasks.

27. Moderation is for teachers to reach agreement about how assessment tasks will be administered to students.

28. Our school needs more help with moderation.

29. Moderation is providing me with professional learning opportunities.

PTO
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>30. Moderation is an accountability exercise.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>31. Moderation meetings are stressful.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>32. Moderation meetings are not useful.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>33. Moderation is helping me to clarify my understanding of the National Standards for writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>34. Moderation is helping me to clarify my understanding of the Literacy Learning Progressions for writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>35. Moderation is helping me to clarify my understanding of my school’s achievement expectations for writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>36. Moderation is making me feel more confident about making dependable judgements of student achievement in writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>37. Moderation is helping me to improve the way I teach writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>38. Moderation is helping me identify my students’ next learning steps for writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>39. Moderation is helping me to improve my classroom assessment in writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>40. I feel comfortable querying/discussing the judgements and ideas that my colleagues express during moderation sessions within my syndicate.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>41. I feel comfortable querying/discussing the judgements and ideas that my colleagues express during inter-syndicate moderation sessions.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>42. I feel comfortable querying/discussing the judgements and ideas that my colleagues express during full school moderation sessions.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>43. I feel comfortable about my colleagues querying/discussing the judgements and ideas that I express during moderation sessions within my syndicate.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>44. I feel comfortable about my colleagues querying/discussing the judgements and ideas that I express during inter-syndicate moderation sessions.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

PTO

<table>
<thead>
<tr>
<th>Question</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>45. I feel comfortable about my colleagues querying/discussing the judgements and ideas that I express during <strong>full school</strong> moderation sessions.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46. I have the opportunity to contribute my ideas and thoughts during moderation sessions within my <strong>syndicate</strong>.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>47. I have the opportunity to contribute my ideas and thoughts during <strong>inter-syndicate</strong> moderation sessions.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>48. I have the opportunity to contribute my ideas and thoughts during <strong>full school</strong> moderation sessions.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

49. Please provide details about any professional learning contracts or initiatives that you have participated in during the last **five years** that have **included a focus on assessment**.  
   Examples include, **Assess to Learn (ATOL), The Literacy Professional Development Project (LPDP), National Monitoring (NEMP or NMSA)**.

50. In general, what do you think have been the benefits of your school’s moderation processes?

51. Do you have any suggestions for how your school could improve its moderation processes?

**THANK YOU FOR TAKING THE TIME TO COMPLETE THIS QUESTIONNAIRE.**
Appendix L. Excerpt from transcription memo

Memo: Transcription of meeting 1 at Central Heights School
- Look for other references to elusive or hard-to-define writing qualities. This came to the teachers’ attention during this meeting as they tried to develop those criteria that encapsulated the notion of “personal voice.”
- During subsequent readings of the data, explore the processes that were used within this session (and others) for clarifying a criterion so that it is easier to measure.
- Explore teacher knowledge/familiarity with the curriculum and other centrally developed documents.
- Review the New Zealand curriculum English exemplar matrices (see Central Heights School meetings folder Meeting 1, Matrices a & b).
- Take note of the extent to which teachers consult/consider the National Standards exemplars (that accompany the achievement descriptors). Because teachers are directed to look at the Literacy Learning Progressions (which do not have exemplars), I think some teachers might rely on the descriptors alone.

Memo: Transcription of meeting 2 at Central Heights School
- Explore the relationship between this school’s moderation activities and their involvement in other initiatives. Consider their involvement in Ministry of Education initiatives (such as the Accelerated Learning in Literacy initiative) and the teacher professional learning opportunities that they access during the year (for example, e-asTTle, Jill Eggleton and Barbara Watson courses).
- Think about the ways in which schools are using the professional development opportunities that are available to them to inform their moderation processes.
- Try to find out more about existing assessment processes/practices that are referred to during observed meetings, for example, “Now, not the usual self-assessment, teacher assessment, peer assessment thing, I just mean something that…” (P. 24, L. 6–8).
- Look at/for references about providing feedback to kids – sharing their Success Criteria with kids.
Appendix M. Excerpt from a chronology

Meeting: 3 Greenville School
Meeting Type: Middle syndicate meeting
Date: 15 May, 2012, 3:15 pm
Location: Middle syndicate meeting room
Present: Susan, Sophia, Phillip

1. Initiating moderation session
   - Outlining the purpose of a meeting (Susan: P. 1, L. 10–12)
   - Referring to past assessment practice (P. 1, L. 27–34)
     *Phillip: I think the way that the Board’s assessment of recounts was done was just*
     *Susan: right, horrendous.*
     *Phillip: dreadful.*
   - Voicing a personal standard/expectation (PP. 1–2, L. 40–)
   - Outlining/reiterating the purpose of a meeting (P. 3, L. 5–15)
   - Valuing or promoting consistency (P. 3, L. 42)
   - Involving students (P. 3, L. 42)
     *Because if they [our school-specific criteria] are consistent, they’ll [the students] find it a lot easier to use them (Sophia: P. 3, L. 42).*

2. Drafting year 2 criteria
   - Outlining the purpose of a meeting (Susan: P. 5, L. 6)
   - Determining how to develop school-specific criteria (P. 5, L. 12–17)
     *• Quantifying/limiting criteria: Devising a rationale for drafting criteria*
     *Sophia: So we’re not putting everything, so we don’t have [referring to the Literacy Learning Progressions]*
     *Susan: No, just a wee feel of, some of them [the criteria] are for us aren’t they, really. (P. 5, L. 12–17)*
     *• Referring to resources (P. 5, L. 14–20)*
     *• Proposing a criterion (Phillip: P. 5, L. 20)*
     *• Questioning or clarifying a criterion (Sophia: L. 24–5)*
   - Voicing a personal standard/expectation: Drawing on personal experience to define/justify school standards
     *I would take this one, “Using their visual memory to spell personal vocabulary” (LLP, p. 13 After 2 years), and I’d actually put that as a “should.” ’Cause when they really get going, some of them aren’t going to do it. . . . It should be a “must,” but it’s probably not going to be (P. 6, L. 21–9).*
     *• Referring to resources*
   - Clarifying a criterion (PP. 6–7)
   - Sharing teaching ideas (P. 7, L. 50–1)
   - Referring to a centrally developed resource: LLP (P. 8, L. 9–11)
   - Drafting a criterion: Use a variety of nouns and verbs (PP. 8–10, L. 32)
     *• Referring to a centrally developed resource: LLP (P. 8, L. 32–3)*
     *• Proposing a criterion [wording] (P. 8, L. 48)*
     *• Questioning a criterion [wording] (P. 8, L. 51)*
     *• Referring to a centrally developed resource: LLP (P. 9, L. 7)*
   - Drafting a criterion: Letter formation
• Referring to a centrally developed resource (P. 11, L. 1–2)
• Proposing a criterion [wording] (L. 8)
• Confirming a criterion wording (L. 11)
  o Drafting a criterion: Simple & compound sentences
    • Referring to a centrally developed resource (P. 11, L. 16–7)
    • Drawing on personal experience to justify expectations
      I would put you “could,” there . . . use different sentence
      beginnings (Phillip: P. 11, L. 23).
    • Questioning a proposed criterion wording: Suggesting an
      alternative wording (P. 11, L. 26)
    • Drawing on personal experience to define/justify expectations
      (I would expect mine to be using them as a “should.” They should
      not be doing “and then, and then and then” (Sophia: P. 11, L. 35–6).
      Otherwise they are just not extending themselves at all (Sophia: P.
      12, L. 8).
    • Developing knowledge of school-wide writing progression: teacher
      learning
      Year 2 is the year level that I have got the least experience with
      (Phillip: P. 11, L. 39).
    • Confirming a criterion (P. 11, L. 29 & 51)
  o Drafting a criterion: Simple conjunctions [P. 12, L. 14]
    • Proposing a criterion wording (L. 14)
    • Referring to a centrally developed resource (L. 14 & 24–5)
      • Involving students
        To join two ideas. Keep it simple for them [the kids] (Sophia: P.
        13, L. 8)
  o Drafting a criterion: verb subject & noun pronoun agreement [P. 13, L. 13]
    • Referring to a centrally developed resource (L. 23)
    • Involving students
      It [the Literacy Learning Progressions] says, “with subject verb
      agreement and noun pronoun agreement” but how the heck do you
      put that in kid speak? (Phillip: P. 13, L. 23–4).
    • Determining how to develop criteria: Specifying teacher only
      criteria (P. 14, L. 10–13)
      Susan: Maybe that is something we …
      Sophia: Have to just help them with …
  o Drafting a criterion: Sentence endings [P. 14, L. 25–]
    • Referring to a centrally developed resource (L. 25–6)
    • Sharing teaching ideas: “Full stopies” (Phillip: L. 47–50)
  o Drafting a criterion: Capitalisation [P. 15, L. 9–]
    • Referring to a centrally developed resource (L. 9–10)
  o Cross-checking with a centrally developed resource (P. 16, L. 15)
  o Drafting a criterion: Spelling rules [P. 16, L. 20]
    • Identifying a gap in criterion coverage (L. 20)
    • Referring to a centrally developed resource (L. 29–32)
Appendix N. Initial codes

- **Utilising textual resources**
  - Using or referring to a centrally developed resource
    - Cross-checking, comparing or aligning school criteria with a centrally developed resource
    - Referring to previous engagement with a centrally developed resource
    - Critiquing a centrally developed resource
      - Voicing frustration with a centrally developed resource
    - Praising a centrally developed resource
    - Promoting the use of centrally developed resources
  - Using or referring to other textual materials
    - Promoting the use of other resources
      - Distributing copies of other resources
    - Critiquing other resources
  - Using or referring to school-specific resources/materials

- **Utilising expertise**
  - Utilising internal expertise
    - Genre expertise
    - Year-level expertise
      - Expressing personal standards
    - Assessment expertise
  - Utilising external expertise
    - Feeding back to colleagues following a relevant course/workshop
    - Referring to relevant professional development (PD) opportunities/experiences
      - Informing colleagues about upcoming PD opportunities
      - Referring to NEMP experience
      - Referring to e-asTTle course
      - Referring to international PD
      - Participating in AToL
      - Attending general National Standards workshops
      - Leadership and Assessment PD
      - Referring to Accelerated Learning in Literacy (ALL) or Accelerated Learning in Mathematics (ALiM)

- **Developing & reviewing school-specific resources**
  - Determining how to develop/review school-specific criteria/exemplars
    - Discussing how to organise/categorise criteria
      - Quantifying/limiting criteria
      - Specifying teacher-only criteria
    - Proposing/questioning/clarifying/revising or affirming a criteria-development process
    - Providing feedback about a criteria-development process
    - Setting norms for participating in criteria development
    - Valuing or promoting specificity during criteria development
    - Discussing how to share/record/review/trial criteria
- Determining a rationale for setting standards/expectations
- Discussing exemplar development

- Drafting/revising a criterion: Development phase (DP) or judgement-making phase (JP)
  - Reading existing criteria aloud
  - Proposing a criterion
  - Revising a criterion
    ⇒ Questioning/clarifying a criterion
      o Clarifying how to define/measure a criterion
      o Discussing the categorisation of a criterion [e.g., surface/deep]
    ⇒ Advocating for or justifying a criterion
    ⇒ Suggesting an alternative criterion
  - Removing a criterion
  - Affirming/Confirming a criterion
  - Questioning coverage or adequacy of school-specific criteria
    ⇒ Identifying a gap in criterion coverage
  - Referring to “trialling” criteria

- Developing school-specific exemplars
  - Creating an exemplar folder
  - Inviting teachers to share photocopied student work to contribute to exemplar folder

  o Developing & reviewing assessment tasks or tools
    - Determining how to develop an assessment task
      - Proposing an idea or giving feedback about the process for developing assessment tasks
    - Developing an assessment task
      - Aligning assessment task with teaching/assessment focus
      - Attempting to improve consistency of a common assessment
      - Determining assessment conditions
        ⇒ Discussing/determining task administration
    - Critiquing/Identifying a problem with a teacher developed assessment task:
      ⇒ During development
      ⇒ After administration
        o Sufficiency of information
        o Assessment conditions
        o “Opportunity” issues
    - Appraising or revising a standardised assessment tool or task
      ⇒ (E.g., NEMP, e-asTTle task, including rubric)
      - Revising or adapting a standardised assessment tool/task
        ⇒ Discussing or clarifying a criterion for a standardised assessment task
      - Critiquing a standardised assessment tool or task
      - Discussing/clarifying assessment conditions

  o Reviewing and evaluating student writing
    - Determining how to review student writing
      - Proposing a process for reviewing student writing
      - Providing feedback about the process of reviewing student writing
- Identifying the aural and visual dimensions or requirements associated with moderation
- Seeking visual access to a writing sample
- Identifying a process for engaging with writing
- Determining when to share/read criteria aloud while reviewing writing
- Valuing discussion of writing qualities over comparison of judgements
- Acknowledging using gut feelings/instincts to inform judgement making
- Identifying/Clarifying criteria used to inform judgement making

- Discussing pre-moderation or sample selection processes
  - Explaining a system for withholding student identity during moderation
  - Explaining a rationale for sample selection
  - Justifying or defending sample selection

- Initiating & inviting the sharing of writing samples
- Explaining or clarifying the assessment task or context
  - Providing details about teaching associated with writing sample
  - Clarifying assessment context/task
  - Explaining/describing conditions of assessment

- Reviewing student writing in relation to agreed criteria
- Reviewing student writing not clearly in relation agreed criteria
- Providing information about a student prior to judgement making
  - Revealing student identity
  - Specifying or clarifying age/year level/time at school
  - Sharing school-entry knowledge/ability
  - Sharing “other” information about a student
  - Identifying student through writing style and voice
  - Previewing writing quality: Providing information before sharing/reviewing

- Providing information about a student after judgement making
  - Revealing/establishing child’s identity

- Declining or attempting to decline invitation to share a writing sample
- Overlooking or omitting an evaluation of a student’s writing
- Comparing writing with an earlier sample
- Determining a judgement-making process
  - Determining how to record judgements
  - Determining how to make OTJs
  - Using gut feelings to inform judgement making
  - Proposing a judgement-making process
  - Providing feedback about a judgement-making process

- Inviting a judgement
  - Inviting a judgement after writing shared & reviewed
  - Inviting a judgement after writing shared but not reviewed

- Sharing a judgement
  - Sharing a judgement after writing shared & reviewed
  - Sharing a judgement after writing shared but not reviewed
  - Sharing a judgement before writing shared
  - Attempting to share a judgement
  - Offering a tentative judgement
• Praising student writing in general terms
• Comparing judgements
  ▪ Comparing writing samples to clarify a judgement
• Avoiding sharing a judgement
• Using National Standards reporting terminology to describe achievement
• Using other reporting terminology to describe achievement
  ▪ Using school-specific reporting terminology to describe achievement
    ⇒ Within; one, two, or three triangles
  ▪ Using e-asTTle rubric reporting terminology to describe achievement
  ▪ Using e-asTTle reporting terminology to describe achievement
  ▪ Using curriculum-level reporting terminology to describe achievement
• Identifying or acknowledging a borderline judgement
  ▪ Using a dual or mixed rating to describe achievement
• Seeking justification of a judgement
• Justifying a judgement
  ▪ Referring to agreed criteria to justify/support a judgement
  ▪ Referring to other criteria or values to justify a judgement
  ▪ Justifying a judgement before sharing & reviewing writing
  ▪ Explaining a judgement
• Clarifying or questioning a judgement
  ▪ Offering an alternative judgement
• Seeking consensus or agreement about a judgement
  ▪ Attempting to confirm a judgement
• Adjusting a judgement
• Confirming or affirming a judgement
  ▪ Achieving consistency or consensus
• Acknowledging a misjudgement or a judgement error
• Recording a judgement
• Offering limited explanation for judgement
• Deviating from emerging judgement-making process
  ▪ Classroom teacher offering judgement first
• Being defensive during judgement making

○ Developing & reviewing processes for documenting & reporting achievement
  • Explaining a process for assessing, documenting, reporting, & tracking achievement
  • Discussing reporting terminology
  • Defining National Standards reporting bands: above, at, below, well below
    ▪ Educating parents about the National Standards reporting bands

○ Leading moderation
  • Identifying a purpose or rationale for moderation
  • Identifying a need [for moderation]
  • Outlining or reiterating the purpose of a moderation meeting
  • Outlining future plans or next steps for moderation
  • Promoting moderation practices
    ▪ Outlining/proposing/confirming a moderation approach
    ▪ Explaining an aspect/feature of a moderation approach
• Defending/justifying a moderation approach/process
  • Setting norms
  • Scheduling & making time for moderation
  • Accessing & promoting resource use
  • Initiating a review of moderation processes
    ▪ Inviting/seeking input/feedback from colleagues
  • Managing colleagues’ concerns and anxieties
    ▪ Acknowledging teacher anxiety associated with moderation
    ▪ Reassuring colleagues
    ▪ Reminding teachers about alignment of school criteria with National Standards
    ▪ Proposing a solution
  • Acknowledging or praising colleagues
  • Aligning practice
    ▪ Promoting developing knowledge of school-wide writing progression
    ▪ Aligning school writing/moderation focus with enrichment (ALL) programme focus
    ▪ Managing/streamlining teacher workloads
  • Refocusing or reorientating colleagues
  • Acknowledging the difficulty of moderation
  • Participating principal
    ▪ Affirming/valuing practice: Principal
    ▪ Thanking lead teacher: Principal
    ▪ Acknowledging colleagues’ work: Principal
  • Leading from the outside
    ▪ Non-participating principal
  • Valuing or appreciating leadership

○ Involving students
  • Employing “kid speak”
  • Sharing criteria with students
  • Promoting self-assessment
  • Promoting sharing student writing with other classes
  • Developing students’ assessment capability

○ Strengthening teaching
  • Linking moderation and & student learning
    ▪ Linking moderation & criteria development with teaching
  • Identifying next steps
    ▪ Inviting next steps
    ▪ Identifying, discussing, or recording next steps when prompted
    ▪ Identifying, discussing, or recording next steps unprompted
    ▪ Identifying students that require additional support
  • Facilitating early intervention
  • Aligning school-wide writing programmes
    ▪ Aligning writing & moderation focus with enrichment (ALL) programme focus
  • Linking moderation with teaching as inquiry
Voicing concern or anxiety
- Threat of mandatory standardised testing
- Spending too much time assessing
- Oversimplifying criteria & limiting brighter children
- Reporting students as “below” their year level
- The accuracy of previous reporting
- Inter-rater reliability

Critiquing the current assessment climate
- Identifying problems with the implementation of the National Standards
  - Voicing frustration about the implementation of the National Standards
- Critiquing the National Standards document or policy
  - Note that this may be a positive or negative critique
- Referring to accountability

Identifying, acknowledging, revealing, or resolving a challenge
- Identifying an organisational challenge
  - Determining how to “do” moderation
  - Struggling to establish school-wide moderation systems
  - Emerging disparity in marking between syndicates
  - Emerging disparity in moderation processes between syndicates
    - Referring variably to drafts or published versions of student writing
  - Developing school-wide, common writing terms
  - Making time for moderation
  - Establishing inter-school moderation processes
- Identifying or revealing a leadership challenge
  - Expressing uncertainty
  - Lacking necessary support
- Identifying a challenge relating to criteria development & use
  - Developing shared understandings
  - Differentiating between year 5 & 6 achievement during criteria development
  - Defining an elusive criterion
  - Emerging tension between teaching & assessment goals
    - Discussing the inclusion of a planning criterion
  - Acknowledging the challenges associated with measuring achievement using criteria
  - Developing common understanding of the National Standards
- Identifying a challenge relating to task development or use
  - Comparing responses to different assessment tasks
- Identifying a challenge relating to judgement making & reporting
  - Evaluating students in years 1–3
  - Making interim or mid-year judgements
    - Anticipating or forecasting progress
  - Noting and reconciling differences between draft & published versions of a writing sample
  - Comparing responses to different assessment tasks
  - Judging borderline students
- Judging students whose achievement sits outside your teaching year level[s]
- Maintaining an open-minded approach during judgement making
- Acknowledging the effect of opportunity issues on performance
- Viewing writing judgements as subjective
- Accessing textual resources
- Coping with multiple textual resources
- Accessing external expertise & support
- Resolving a challenge
- Identifying or revealing a challenge relating to converting results or levels
  - Converting NEMP scores to National Standards scores/reporting bands
  - Aligning NEMP tasks with the National Standards
- Identifying a challenge relating to reporting
  - Leap between “after 1–3 years at school” to “at the end of year 4”
- Referring to past assessment practice
  - Developing & using criteria, checklists, or exemplars
  - Moderating, benchmarking, or comparing judgements
  - Developing and/or administering common assessment tasks
  - Reporting practices
  - Describing early National Standards practices
- Shaping dependability
  - Valuing improving reliability & validity
  - Valuing or promoting consistency
    - Standardising practice
    - Promoting school-wide alignment
    - Developing a common language
    - Valuing or prioritising consistency in judgement making
      ⇒ Valuing consistency over sharing judgement-making rationale
    - Improving consistency or accuracy
    - Modifying practice to ensure alignment
      ⇒ Modifying & adjusting practice
    - Reaching agreement
    - Improving inter-school consistency
    - Questioning inter-school consistency
- Identifying & addressing potential sources of bias
  - Identifying potential sources of teacher or assessor bias
    - Acknowledging that other knowledge of a student can affect judgement making
      ⇒ Acknowledging the effect of assessing an unknown student: e-asTTle session
    - Acknowledging the effect of assessing an unknown student during an e-asTTle session
  - Discussing processes that may help reduce the incidence of teacher or assessor bias
- Reading writing aloud to reduce/limit the focus on surface features
- Identifying & examining the possible effect of gender

- Acknowledging the role that criteria play in ensuring objectivity
- Identifying and/or addressing potential task biases
- Identifying opportunity issues

- **Converting results**
  - Converting curriculum levels or e-asTTle levels to National Standards year levels/reporting bands
  - Sharing a conversion rationale
  - Devising a conversion rationale
  - Referring to a conversion rationale

- **Introducing or correcting an error**
  - Introducing an error while developing & reviewing school-specific criteria & exemplars
  - Correcting an error while developing & reviewing school-specific criteria & exemplars
  - Introducing an error while reviewing student writing
  - Correcting an error while reviewing student writing
  - Introducing an error while judging student writing
  - Correcting an error while judging student writing

- **Articulating personal standards**
  - Voicing a personal standard/expectation
    - Valuing a specific criterion
    - Voicing a personal standard
    - Profiling ability at a given level
    - Drawing on personal experience to define/justify expectations
    - Justifying a personal standard or expectation
    - Cross-checking criteria with teacher expectations
    - Comparing expected or typical performance with actual performance

- **Improving teachers’ assessment capability**
  - Improving judgement-making capability
    - Improving accuracy/consistency of judgement making
    - Strengthening confidence [about judgement making]
      ⇒ Allaying fears
  - Improving task administration
  - Improving the quality and specificity of reporting
  - Acknowledging teacher learning
    - Acknowledging the role of teacher learning in relation to moderation
    - Acknowledging moderation sessions as an opportunity to share prior teacher learning
    - Developing knowledge of school-wide writing progression
    - Acknowledging the iterative/ongoing nature of moderation
    - Developing knowledge of centrally developed resources through criteria development
- **Sharing teaching ideas**
  - *Sharing details about recent changes to writing instruction*
  - *Discussing writing courses or pedagogical approaches*
  - *Acknowledging the link between teaching and achievement*

- **Owning or valuing moderation**
  - Articulating a purpose or rationale for moderation
  - Acknowledging the benefits associated with moderation
  - Valuing school-specific processes or products
  - Valuing full-staff meeting sessions
    - Valuing full-staff meeting sessions for developing knowledge of the school-wide progression
  - Valuing syndicate meeting sessions
  - Valuing e-asTTle moderation session
  - Valuing criteria-development sessions
    - Valuing syndicate-level criteria-development sessions
  - Valuing judgement-making sessions
    - Valuing paired moderation meeting
  - Valuing opportunities to collaborate
    - *Reducing isolation*
    - *Working together*
    - *Opening up practice*
    - *Strengthening or building relationships*

- **Acknowledging/involving researcher**
  - Acknowledging or involving researcher
  - Participating researcher
  - Revealing researcher/observer effect
  - Clarifying researchers’ “needs”
Appendix O. Central Heights School moderation plan

**Literacy PD for 2012**

- Teacher Only Day analyse National Standards Data and discuss the provisional target for the year.
- Read over the National Standards for each teacher’s only main level.
- Week 3 Syndicates plan success criteria for personal recount for the levels they teach. Teachers refer to more useful criteria for achievement in the Literacy Progressions for writing.
- Week 4 staff meets to formulate a school wide success criteria for personal recounts.
- Next 6 weeks, staff teach the writing of personal recount. Use exemplars where appropriate.
- Week 10, staff have a moderation staff meeting.
- Term 2, 3, and 4 the same procedures will continue but genre is not yet decided.
- Syndicates will moderate their own children’s writing during each term before the whole staff meeting at the end of each term.
Appendix P. NEMP assessment task: Torch


### Trend Task: Torch

<table>
<thead>
<tr>
<th>Approach</th>
<th>Station</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>Writing instructions</td>
</tr>
<tr>
<td>Resources</td>
<td>Assembled torch without batteries, 2 batteries</td>
</tr>
</tbody>
</table>

#### Questions / Instructions:

- Check that you have a torch [ ] and two batteries [ ]. If not, tell the teacher.
- Put the batteries in the torch to make it go. If the light does not go, tell the teacher.
- Write instructions so that a young person would know how to put the batteries in the torch to make it go.

#### Instructions included:

- taking the top off
- putting the two batteries in
- awareness that battery orientation matters
- getting the batteries in the same orientation
- conventional orientation of batteries (positive towards the bulb or negative towards the spring)
- putting the top back on
- switching the torch on

### % Response

<table>
<thead>
<tr>
<th></th>
<th>Year 4</th>
<th>Year 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Mid Range

#### Commentary

Boys are often thought to be more interested than girls in how to make technological devices like torches work, but girls performed better than boys on this task. Thirty-five percent more year 8 than year 4 students gave instructions that achieved a score of six or seven. There was no meaningful change in performance between 2002 and 2006.
## Appendix Q. Sample of a Tracking Sheet

### Writing Expectations End Year 3 Tracking Sheet

<table>
<thead>
<tr>
<th>Group</th>
<th>Purpose and Audience</th>
<th>Punctuation</th>
<th>Phonics and Spelling</th>
<th>Sentence Structure</th>
<th>Review and Rereading</th>
<th>Use speech marks when needed</th>
<th>Spell some Essential Lists 5 &amp; 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Use planning to organise ideas</td>
<td>*</td>
<td>Correct use of full stops</td>
<td>Use simple and compound sentences</td>
<td></td>
<td>Use correct form when questioning</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Use correct text structure</td>
<td></td>
<td>Correctly use question marks</td>
<td>Use different sentence beginnings</td>
<td></td>
<td>Use correct form when spelling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Independently correct texts</td>
<td></td>
<td>Correctly use full stops</td>
<td>Use a variety of nouns, adjectives and verbs</td>
<td></td>
<td>Use correct form when punctuation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Understand the purpose of their writing</td>
<td></td>
<td>Use simple and compound sentences</td>
<td>Use language features eg. Alliteration</td>
<td></td>
<td>Proofread to check grammar and make some corrections</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Appropriate choice of language and text form</td>
<td></td>
<td>Use correct form when spelling</td>
<td>Write letters correctly, legibly and fluently</td>
<td></td>
<td>Proofread to check punctuation and make some corrections</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Give parents feedback on their writing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Should</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Should</td>
<td>Spell some Essential Lists 5 &amp; 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Should</td>
<td>Use correct form when spelling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Should</td>
<td>Use correct form when punctuation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Should</td>
<td>Use correct form when spelling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Should</td>
<td>Use correct form when punctuation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Should</td>
<td>Use correct form when punctuation</td>
</tr>
</tbody>
</table>

Notes: * Not all children's Must, Should, Could apply.
Appendix R. Previously published manuscript

This is an Accepted Manuscript of an article published by Taylor & Francis in Assessment in Education: Principles Policy & Practice, 2013, available online: http://dx.doi.org/10.1080/0969594X.2012.696241

Moderating New Zealand’s National Standards: teacher learning and assessment outcomes

Esther Smaill*

Educational Assessment Research Unit, University of Otago, Dunedin, New Zealand

(Received 8 February 2012; final version received 19 May 2012)

Across the globe, standards-based assessment systems are increasingly promoted as a means of improving student outcomes and fulfilling accountability requirements. Within such systems, social moderation is presented as a mechanism for improving the dependability and utility of assessment information. Research emphasises that social moderation processes provide professional learning opportunities; yet this learning tends to be perceived as a by-product rather than a goal of moderation. Situated within the context of New Zealand’s recently implemented National Standards, this article reviews the literature and presents an alternative configuration of the role of social moderation. It foregrounds the professional learning that will be required if these moderation processes are to improve the dependability of National Standards assessment information and contends that reconfiguring teacher professional learning as a goal rather than a by-product of moderation should not only improve dependability and strengthen teacher assessment capability but also assist with lifting student outcomes.

Keywords: social moderation; teacher professional learning; standards; teacher judgement making; teacher summative assessment

Introduction

Globally, standards-based assessment systems are increasingly seen as a means of improving student outcomes and meeting accountability requirements (Klenowski and Wyatt-Smith 2010a). In this standards-driven climate, the use of teacher summative assessment, as opposed to national testing, is gaining salience (Harlen 2005). Where teacher assessment is used for summative purposes, moderation systems are often introduced to ensure the dependability of assessment information (Maxwell 2010). Although there is widespread recognition that engaging in moderation can provide teachers with professional learning opportunities (Gardner et al. 2010; Gilmore 2008; Gipps 1994; Harlen 1994a, 2005; Hipkins 2010a; Hipkins and Robertson 2011; Klenowski and Wyatt-Smith 2010a; Linn 1993; Malone, Long, and DeLucchi 2004; Maxwell 2002a, 2010; Radnor and Shaw 1995; Van Krayenhondt et al. 1999), this learning has tended to be presented as a by-product rather than a goal of moderation. In response to burgeoning acknowledgement that teacher learning plays a critical role in moderation (Connolly, Klenowski, and Wyatt-Smith 2011; Klenowski and Adie 2009; Klenowski and Wyatt-Smith 2010b; Wyatt-Smith, Klenowski, and Gunn 2010), this article reviews the literature and examines the

*Email: esther.smaill@otago.ac.nz

© 2013 Taylor & Francis
professional learning that is required to ensure the efficacy of an approach termed *social moderation*. Located within the context of New Zealand’s recently introduced National Standards assessment system, this article foregrounds the complexity and extent of the teacher learning that social moderation necessitates. Specifically, it shows that in addition to using moderation meetings to develop shared understandings of the new National Standards, teachers will also need to use these sessions to strengthen their judgement making and evidence collection.

An examination of comparable moderation systems in both Australia and the United Kingdom reveals that, although participation in social moderation can facilitate professional learning, thereby strengthening teacher assessment capability, it can also result in missed learning opportunities and unintended learning. Such unexpected and unharnessed learning can threaten the dependability of assessment information and subvert the educative benefits of social moderation. But much of the variability in the teacher learning opportunities associated with social moderation processes can be explained when analysed in relation to evidence-based research detailing the characteristics of effective professional learning. Reconfiguring teacher professional learning as a goal, rather than a by-product, of social moderation and utilising evidence-based research on effective professional learning to inform and resource these processes should therefore facilitate improvements in both dependability and teacher assessment capability. Moreover, because professional learning opportunities that include a focus on strengthening teacher assessment capability have been linked with improved student outcomes (Timperley et al. 2007), ensuring the efficacy of social moderation processes could assist with lifting student achievement.

**New Zealand context**

In 2010 the New Zealand Government introduced a system of achievement standards, called the National Standards, for English-medium schools with students in years 1–8 (Ministry of Education 2009a). The following year, comparable standards were introduced for Māori-medium schools catering for students in the same age group (Ministry of Education 2010a). These standards, which are expected to enable students to meet the demands of the New Zealand Curriculum (Ministry of Education 2009a), specify yearly achievement expectations in reading, writing, and mathematics. The National Standards consist of descriptors, exemplars of student work and assessment tasks linked to school year levels (Ministry of Education 2010b). They are a form of what Maxwell describes as standards as quality benchmarks, or the specification of ‘expected practice or achievement’ (2002b, 2).

Mirroring a global trend in education (Klenowski and Wyatt-Smith 2010a), New Zealand’s recently introduced National Standards are expected not only to improve educational outcomes (Ministry of Education 2009a, 2009b) but also, according to the education manifesto of the governing National Party, to strengthen accountability (New Zealand National Party 2011). Although teachers are required to assess all students against this set of common standards, the approach to gathering assessment information is not tightly specified, and there is no national test (Ministry of Education 2011). In fact, because the use of common assessment tasks might not provide all students with the opportunity to perform optimally (Harlen 1994b; Ministry of Education 2010c), teachers are encouraged to tailor their selection of both self-generated and centrally developed assessments to the needs of individual students.
(Ministry of Education 2010d). Teachers then use each student’s unique evidence base to inform a qualitative judgement of that student’s achievement in relation to the National Standards (Ministry of Education 2010d). These summative judgements, called ‘overall teacher judgements’, must be made twice yearly for each student in relation to the standards specified for reading, writing and mathematics (Ministry of Education 2010e). A comparable approach has been adopted in Wales (Estyn 2010), where a system of teacher assessment has replaced the statutory testing of 11 and 14 year olds. In contrast, national testing has never been a feature of New Zealand’s year 1–8 school system.

Prior to 2010, and since the introduction of self-management in 1988, New Zealand’s primary and intermediate schools (henceforth referred to as primary schools) had considerable freedom to choose how to approach assessment and reporting with respect to the National Education Guidelines (Absalom et al. 2009; Ministry of Education 2010b). But amendments made to the National Administration Guidelines since the introduction of National Standards have resulted in changes for educators that signal a major shift in New Zealand’s assessment and reporting policy. Schools are now required to report annually on, ‘the numbers and proportions of students achieving at, above, below or well below the standards’ (Ministry of Education 2010f). In this new environment of increased accountability, teachers are also expected to participate in intra- and ‘where appropriate’, inter-school moderation processes as part of system-level efforts to improve the dependability of student assessment information (Ministry of Education 2010g, 2).

The need to moderate arises in part because of the nature and quality of achievement standards. New Zealand’s National Standards are a form of what Sadler (1987) describes as verbal descriptions or qualitative rubrics. Because such standards are specified in linguistic terms, their meanings are ‘always to some degree vague or fuzzy’ (Sadler 1987, 202), and it is widely recognised that they require interpretation (Connolly, Klenowski, and Wyatt-Smith 2011; Klenowski and Adie 2009; Klenowski and Wyatt-Smith 2010a, 2010b; Maxwell 2002a, 2002b, 2009, 2010; O’Donovan, Price, and Rust 2004; Sadler 1987, 2009; Wyatt-Smith, Klenowski, and Gunn 2010). Given that an assessor’s interpretation of a standard is shaped by his or her unique evaluative experience (Sadler 1987), and indeed the assessment context (Sadler 2009), these interpretations inevitably differ. Because of this, a number of conditions must be fulfilled to help assessors clarify their interpretations of standards and develop common understandings of their meanings (Klenowski and Wyatt-Smith 2010a; O’Donovan, Price, and Rust 2004; Sadler 1987). Assessors first require exemplars or concrete examples demonstrating the specified levels of achievement. Yet, though necessary, exemplars are not on their own sufficient to ensure that assessors develop shared understandings of achievement standards. Crucially, assessors also need opportunities to participate in the moderation processes that are the focus of this article.

Moderation processes are well established in New Zealand’s secondary schools (Hipkins 2010b; Hipkins and Robertson 2011) where the outcomes of standards-based National Certificate of Educational Achievement (NCEA) assessments influence students’ tertiary and employment opportunities (Ministry of Education 2011). Due to the high stakes associated with these assessments and the importance of ensuring their integrity, NCEA moderation involves both internal and external processes (Ministry of Education 2011). These include the submission of assessment materials for pre-moderation, school-based moderation or co-marking of student
work, and the submission of marked assessments for external moderation. Similar approaches to moderation are used in the senior phase of schooling in Queensland, Australia (Maxwell 2010; Queensland Studies Authority 2010), and at the end of Key Stage 3 in Wales (Estyn 2010).

New Zealand secondary school teachers took several assessment cycles to clarify their understandings of NCEA standards and reach relative consistency in their judgement making (Hipkins and Robertson 2011). Because moderation is a new requirement for New Zealand’s primary school teachers, a similar period of learning and adjustment is expected in this sector (Hipkins and Robertson 2011). Indeed, the evolution of NCEA moderation processes (Hipkins 2010b) suggests that capacity will most likely need to be built throughout the sector to ensure the efficacy of National Standards moderation. Although it is acknowledged that learning will be required at all levels, the focus of this article is on the teacher professional learning that moderation will necessitate. A close examination of the moderation model specified for National Standards purposes will reveal the full extent of this learning.

Social moderation

Given that New Zealand’s National Standards are associated with relatively lower stakes than NCEA assessments, a somewhat less prescribed moderation system has been adopted at the primary level (Ministry of Education 2011). Instead, New Zealand’s primary school teachers are required to engage in a process that is most commonly described as social moderation (Linn 1993; Maxwell 2010). This approach, which has also been referred to as group moderation (Gipps 1994; Harlen 1994b), involves the comparison and alignment of participants’ judgements of student work in relation to stated standards (Connolly, Klenowski, and Wyatt-Smith 2011; Maxwell 2010). During this participative process, assessors clarify their interpretations of qualitative descriptors and develop shared understandings of their meaning (Harlen 1994b; Maxwell 2002a, 2010). Maxwell (2010, 469) emphasises that social moderation should be framed, ‘as a normal and positive consultative process that values consensus and builds confidence’. Consensus is achieved when assessors reach agreement about the standard awarded to a student’s assessment (Connolly, Klenowski, and Wyatt-Smith 2011). Linn (1993) posits that widening social moderation processes to an inter-school level can both broaden consensus and strengthen public confidence, and New Zealand primary schools have been encouraged to establish inter-school moderation networks (Ministry of Education 2010g).

Social moderation processes can serve a dual function (Harlen 1994b; Maxwell 2002a, 2010), fulfilling what Harlen (1994b) describes as both quality control and quality assurance purposes. Maxwell (2002a, 1) uses the terms ‘moderation for accountability’ and ‘moderation for improvement’ to make a similar distinction. As the terminology implies, quality control or accountability functions (which will henceforth be referred to as quality control) are product oriented and involve the adjustment or verification of assessment outcomes to ensure fairness and comparability (Harlen 1994b; Maxwell 2002a, 2010). In contrast, improvement or quality assurance functions (which will from now on be referred to as improvement functions) are largely process oriented. Moderation for improvement strengthens the capability of participants to conduct appropriate assessments and arrive at consistent and comparable judgements (Harlen 1994b; Maxwell 2002a, 2010). Ideally, the goal is for each of these functions to complement the other (Maxwell 2010).
New Zealand’s National Standards moderation processes are expected to fulfil both quality control and improvement functions. Indeed, moderation for improvement is consistent with the notion of ‘teaching as inquiry’ that is promoted in the New Zealand Curriculum (Ministry of Education 2007, 35). Specifically, these moderation processes can strengthen the ways in which teachers conduct, analyse and respond to student assessments. Yet, although teachers have been encouraged to view National Standards moderation as a ‘learning exercise’ rather than just an opportunity to check the accuracy of their judgements (Ministry of Education 2010c, 53), schools have also been advised that a lack of improvement in the percentage of agreement among teachers’ judgements signals that moderation processes are not achieving the ‘desired outcome’ (Evaluation Associates 2010). Public confidence in such moderation systems, which is typically achieved through the provision of information about the degree of inter-rater agreement (Linn 1993), is generally promoted through a focus on quality control. In accordance with this, New Zealand’s Ministry of Education has stated that National Standards moderation processes are expected to provide, ‘assurance to parents and others that interpretations of students’ achievements are in line with other professionals’ (2010c, 12).

In contrast with New Zealand’s NCEA system, there are no mechanisms in place for externally moderating or verifying teacher-generated assessment tasks or teacher judgements associated with the National Standards. Teachers sometimes view external moderation negatively because it can serve an accountability function (Hopkins 2010b). Yet, given the right conditions, external moderation can also provide teachers with valuable feedback about their understanding of standards and the quality of work associated with different levels of achievement (Hopkins 2010b; Maxwell 2010). Indeed, this kind of access to external expertise is nearly always a feature of effective professional learning opportunities (Timperley et al. 2007). With respect to the National Standards, the lack of such an external moderation process eliminates a feedback loop that might otherwise have both facilitated and informed teacher professional learning.

There are a range of professional learning requirements associated with National Standards moderation, and a detailed examination of just what this moderation is expected to deliver reveals the learning it will necessitate. Specifically, National Standards moderation processes are expected to improve the dependability not only of teacher judgements but also of the assessment evidence that is used to inform and support those judgements (Ministry of Education 2010d). The concept of dependability, which can be expressed as the sum of reliability and validity (James 1998), attempts to recognise the inherent tension that exists between these two properties (Gipps 1994; Harlen 1994b, 2005; James 1998). Because efforts to improve the reliability of an assessment are typically associated with a reduction in validity, and vice versa, what matters in practice is the way that these two concepts are combined (Harlen 1994b, 2005; James 1998). Harlen (2005) has suggested that the nature of this combination should be determined by the purpose of an assessment approach. Because New Zealand’s National Standards system attempts to preserve construct validity through the use of teacher assessment for summative purposes rather than national testing, a dependable evidence base might reasonably be defined as one that achieves optimum reliability while preserving construct validity (Harlen 2005).

For participation in National Standards moderation to improve the dependability of teachers’ evidence bases, teacher moderators will need to have opportunities to
strengthen both the reliability and the construct validity of their assessment evidence. It is generally recognised that reliability is affected by the administration or conditions of assessment (Crooks, Kane, and Cohen 1996; James 1998) and the sufficiency of the information gathered (Crooks, Kane, and Cohen 1996; Darr 2005a; Smith 2003). This indicates that for moderation processes to strengthen the reliability of assessment evidence, teachers will need to use these processes to review not only how assessments are designed and administered to students but also whether sufficient information has been gathered to assess the intended learning. To examine the construct validity of an assessment, teachers require a strong understanding of both the construct in question and how that construct is exhibited (Darr 2005b). They require this same level of understanding to strengthen the construct validity of an assessment base.

Given that understandings of the National Standards are still evolving and common understandings have not yet been achieved, strengthening the construct validity of an assessment or an assessment base necessarily involves teachers using moderation processes to engage with and clarify their interpretations of the stated standards. In addition, since the opportunity students are afforded to demonstrate their knowledge and ability in the specified area is a key feature of validity (Harlen 1994b), teachers will need to use moderation sessions to review whether the assessment tools and tasks that were used provided their students with such opportunities. Where teachers are able to use moderation activities not only to strengthen their assessment capabilities in the ways specified but also to prompt the requisite adjustments in their assessment practice, improvements in classroom assessment and the dependability of teachers’ evidence bases should follow.

Within the National Standards context, when the concept of dependability is applied to teacher judgements rather than assessment evidence, an appropriate level of dependability might be defined as one that achieves optimum inter-rater reliability while preserving the validity of the judgement process. Where validity is expressed as an evaluation of the interpretations and decisions that are made on the basis of assessment evidence (Darr 2005b; Messick 1989), the implication is that National Standards moderation processes are expected to improve not only inter-rater reliability, or levels of teacher agreement around judgement making, but also the quality of the inferences and interpretations on which these judgements are based.

With respect to judgement making, potential threats to dependability can arise from varied interpretations of both standards (Harlen 1994b) and assessment information (Crooks, Kane, and Cohen 1996; Harlen 1994b). Differing interpretations of assessment evidence can occur for a range of reasons and might be attributed to inconsistent or inadequate approaches to the aggregation of assessment information (Crooks, Kane, and Cohen 1996), or the intrusion of irrelevant contextual information (Harlen 1994b), which may include teacher biases (Crooks, Kane, and Cohen 1996). If social moderation is to reduce such variability, teachers will need to use moderation to garner a range of learning opportunities. They will need to use moderation sessions not only to develop shared understandings of the stated standards but also to strengthen the ways in which they evaluate and interpret assessment evidence. Where moderation processes are utilised to build teacher assessment capability in each of these ways, these processes should improve the dependability of assessment information and better equip teachers to identify and respond to student learning needs.
The purpose of explicating the professional learning that will be required if social moderation processes are to fulfil expectations and improve the dependability of National Standards assessment information is to demonstrate the extent and complexity of the work that moderation demands. Expectations that these social moderation processes will improve the dependability not only of teachers’ judgements but also of their assessment evidence are presented despite recent acknowledgement that there has been limited research examining the efficacy of such processes (Klenowski and Adie 2009; Maxwell 2010). Given the teacher learning required for social moderation, the extensive evidence-based literature on effective professional learning (Timperley et al. 2007) may offer a way forward.

The existence of important synergies between social moderation processes and those qualities associated with successful professional learning initiatives further justifies a re-analysis of existing research on social moderation through the literature on professional learning. Timperley et al. (2007) have identified a range of characteristics that promote effective professional learning opportunities. Of these, several key features are commonly evident in social moderation processes. These features may include, but are not limited to, the opportunities that participation in moderation processes can afford teachers to: engage in learning over an extended period; use assessment as a catalyst for further learning; challenge prevailing discourses, such as teacher bias; and, in some cases, work with external experts. Examining the access that teacher moderators have to each of these opportunities allows a re-analysis of the existing research base on social moderation. This re-analysis will begin with an evaluation of the opportunities that moderation processes afford teachers for establishing common understandings of stated standards.

**Shared understandings of standards**

New Zealand’s National Standards were developed rapidly (Hipkins and Robertson 2011) and teachers were provided with few opportunities to be involved in this process (Crooks 2011). Although some professional development accompanied the introduction of these standards, the quality of this has been mixed (Education Review Office 2010; Thomas and Ward 2011), and some schools have had trouble accessing adequate support (Education Review Office 2010). Because of this, participation in moderation will play an especially important role in enabling New Zealand primary school teachers to establish shared understandings of the National Standards. In this regard, recent research conducted in Queensland, Australia, provides some critical insights into how participation in moderation processes might inform New Zealand teachers’ development of shared understandings of their new standards. This research (Klenowski and Adie 2009; Klenowski and Wyatt-Smith 2010b; Wyatt-Smith, Klenowski, and Gunn 2010) examined the social moderation processes associated with the Queensland Curriculum, Assessment and Reporting (QCAR) trial, in which teachers worked collaboratively to moderate Queensland Comparable Assessment Tasks. These centrally devised assessment tasks were developed to assess students in Years 4, 6 and 9 in English, mathematics and science. Like New Zealand, Queensland has a history of externally moderated school-based assessment in the senior phase of schooling, but has only recently implemented standards-referenced moderation processes in Years 1–9 (Klenowski and Adie 2009; Klenowski and Wyatt-Smith 2010b; Wyatt-Smith, Klenowski, and Gunn 2010).
The teacher moderators who participated in the QCAR trial worked in groups and were provided with a range of resource materials to guide their work (Klenowski and Adie 2009; Wyatt-Smith, Klenowski, and Gunn 2010). These included a guide, with stated grade descriptors or standards, and annotated examples of student work corresponding to each achievement level. Because these teachers were working to moderate students’ responses to centrally devised assessment tasks, their use of the stated standards and accompanying resource materials informed their judgement-making processes. Although most groups of teachers attempted to use the resources that had been provided (Klenowski and Adie 2009; Wyatt-Smith, Klenowski, and Gunn 2010), four different approaches to resource use were identified. Notably, some teachers reached agreement about the quality of student assessment evidence with little or no reference to the stated materials that had been provided (Klenowski and Adie 2009; Wyatt-Smith, Klenowski, and Gunn 2010). Instead, such groups of teachers, who were typically relatively experienced, based their decisions on unstated personal standards. Although these standards were not discussed, the teachers assumed that they were held in common (Wyatt-Smith, Klenowski, and Gunn 2010). In general, those teachers who did make reference to the stated resource materials adopted one of three approaches. Some attended almost exclusively to the stated standards (the guide) and made very limited reference to the annotated exemplars. Others paid minimal attention to the stated standards and instead relied on the exemplar material to inform their judgements. Others still used the resources in combination, as had been intended.

Because participation in moderation is expected to enable teachers to develop shared understandings of stated standards, it follows that their use of these standards within moderation sessions must inform the understandings that they achieve. Within the QCAR trial, groups of teachers engaged with the stated standards and the accompanying referents in different ways and in some cases scarcely at all. In many cases, the various resources had ‘the unexpected and unintended impact of competing for teacher attention’ (Wyatt-Smith, Klenowski, and Gunn 2010, 65). This prompted these researchers to emphasise the importance of ensuring conceptual clarity in the presentation of such support materials to teachers. It would appear that a lack of conceptual clarity may have constrained the development of shared understandings of standards and so limited an improvement in the dependability of assessment information. Timperley et al. (2007) have demonstrated that the involvement of external experts can promote effective learning opportunities. In this instance, access to such expertise may have assisted with ensuring conceptual clarity, thereby optimising the dependability of assessment information.

The challenges associated with establishing shared understandings of standards are also alluded to in a 2010 report evaluating the arrangements to assure the consistency of teacher assessment in Wales at Key Stages 2 and 3 (Estyn 2010). Following the Welsh Assembly Government’s discontinuation of statutory end-of-key-stage testing for 11 and 14 year olds in 2005, Wales adopted a system of teacher assessment. Welsh teachers are required to take part in school- and cluster-based standardisation and moderation processes to support effective teacher assessments and ensure that pupils are assessed accurately and consistently through shared understandings of national standards.

But five years after this system was implemented, the 2010 evaluation report concluded that although teachers were becoming more confident about their understanding of the national curriculum (NC) level descriptions, differing interpretations
of standards continued to persist (Estyn 2010). Indeed, some schools raised concerns about the difficulty of ensuring that teachers applied ‘the cluster’s shared understandings of standards’ in their subsequent school-based assessments (Estyn 2010, 4). Interestingly, this report also revealed that in the few local authorities where advisory staff members attended all standardisation and moderation meetings, their presence had, ‘a very positive impact on the consistency of application of NC level descriptions within and across clusters of schools’ (Estyn 2010, 12). This is consistent with evidence-based findings that link effective teacher professional learning outcomes with access to learning opportunities that occur over an extended period of time and involve frequent contact with a provider or external expert (Timperley et al. 2007). This lends further weight to the earlier suggestion that the nature and quality of the resourcing associated with moderation affects the extent to which teacher participants develop shared understandings of standards. In this instance, the provision of sustained resourcing at an adequate level, specifically the presence of external advisors, appears to have had the effect of assisting teacher moderators to develop and apply shared understandings of stated standards.

Assessment evidence
Within the context of New Zealand’s National Standards, social moderation processes are also expected to improve the dependability of the assessment evidence on which teachers base their judgements of student achievement (Ministry of Education 2010d). Although the quality of the assessment information that teachers use for summative purposes varies (Harlen 2005; Mitchell and Poskitt 2010; Thomas and Ward 2011; Wylie and Hodgson 2010), there has been limited research examining how moderation shapes teachers’ collection of summative assessment evidence. But findings from a recent project conducted in the United Kingdom provide some indication that moderation meetings can support teachers to use broader evidence bases (Gardner et al. 2010). The final phase of this project, which was designed to strengthen teacher summative assessment at the end of Key Stage 2, required participating year 6 teachers to gather assessment evidence over a period of two terms and attend two moderation meetings. By the time of the second moderation meeting, teachers were gathering a wider range of assessment evidence, from which it was inferred that these moderation sessions had facilitated teacher professional learning.

A complementary explanation drawing on research on professional learning might place more weight on the role of advisors. During the initial phase of the UK project, local authority advisors worked collaboratively with a smaller group of teacher participants to develop context-relevant resources that were then used throughout the project. Practices like this, in which external experts involve teachers in the creation of meaning, have been shown to be more effective than those that require teachers to mechanistically implement a preferred approach (Timperley et al. 2007). In this case, the expertise of the external advisors could help to account for the level of professional learning that teacher moderators experienced in the final phase of this project.

In Australia, some participants in the QCAR trial described above used a moderation session to critique a centrally devised assessment instrument (Klenowski and Wyatt-Smith 2010b). In this context of engagement and co-construction, a group of teacher moderators sought to mitigate the effects of ‘construct under-representation’
through making adjustments to the marking criteria (Klenowski and Wyatt-Smith 2010b, 30). On this basis, it might be inferred that where teachers use moderation as an opportunity to critically analyse the quality and suitability of assessment tools, and where the insights gleaned from this are not only valid but are also subsequently applied to the development of future assessment tools and tasks, the quality of teacher assessments and the dependability of teacher-generated evidence bases might be strengthened. But these gains cannot be expected without the provision of adequate resourcing. Specifically, teachers will require regular opportunities to take part in well facilitated, externally supported, moderation processes.

Judgement making
Social moderation is also presented as a mechanism that can strengthen and align teachers’ judgement making (Harlen 1994b; Klenowski and Wyatt-Smith 2010a; Linn 1993; Maxwell 2010). Within the National Standards context, it is recognised that the demands associated with making overall judgements of student achievement can present a particular challenge (Education Review Office 2010; Mitchell and Poskitt 2010; Wylie and Hodgen 2010). Teachers are expected to base these overall judgements on a range of evidence from a variety of sources and have been encouraged to tailor assessments to individual student needs (Ministry of Education 2010d). Positioning the student at the centre of learning in this way, which is consistent with the principles espoused in the New Zealand curriculum (Ministry of Education 2007), means that the content of student assessment profiles should differ. In order to make an overall judgement, teachers must aggregate or combine each student’s assessment information and arrive at an ‘on balance’ judgement of their achievement in relation to the National Standards (Ministry of Education 2011, 12). Because there are a number of assessment threats associated with the aggregation process (Crooks, Kane, and Cohen 1996), the implication is that moderation will help to reduce these threats and in so doing, improve the dependability of assessment information. In other words, the expectation is that moderation procedures will help to reduce variability in teacher judgements by enabling teachers to learn or develop not only common but also suitably valid approaches to this aggregation process.

The QCAR trial conducted in Queensland, Australia, required teacher moderators to aggregate score components to obtain an overall task grade for each student (Klenowski and Adie 2009; Klenowski and Wyatt-Smith 2010b; Wyatt-Smith, Klenowski, and Gunn 2010). Because of this, an examination of the various aggregation approaches utilised by these teacher moderators affords some insight into how social moderation can shape, regulate and validate teachers’ aggregation processes. In the QCAR trial, the information provided about how to combine score components to make overall judgements was interpreted in a variety of ways, and teachers adopted a range of aggregation processes. Although some groups of teachers proceeded as was intended and arrived at overall judgements through an on-balance evaluation of the task as a whole (Klenowski and Adie 2009), other groups devised different methods. These included averaging the grades for each assessable element (Klenowski and Adie 2009; Wyatt-Smith, Klenowski, and Gunn 2010), allocating the responsibility for making overall judgements to a senior teacher (Klenowski and Wyatt-Smith 2010b), reaching a decision through valuing either specific aspects of the task or certain assessment criteria over others (Klenowski and Adie 2009), and
creating and then totalling numeric sub-scores for each of the criteria to reach an overall grade (Klenowski and Wyatt-Smith 2010b).

Though participation in moderation undoubtedly provided these teachers with opportunities to learn or develop aggregation processes, not all the methods adopted were suitably valid. For example, where moderation processes involve the averaging of judgements, opportunities for participants to challenge and clarify interpretations and applications of standards are limited (Maxwell 2010). Likewise, obtaining a total score by averaging individual assessable elements limits opportunities for teacher moderators to examine, clarify and thereby validate their aggregation processes. Moreover, for an aggregated or overall score to be most meaningful, the rationale for allocating varying weights to different tasks or task components should reflect their relative importance (Crooks, Kane, and Cohen 1996). In the QCAR trial, limited guidance was provided about how to approach aggregation (Klenowski and Wyatt-Smith 2010b; Wyatt-Smith, Klenowski, and Gunn 2010), and some teacher moderators devised their own rationales for allocating weights to task components. When occurrences such as this result in inappropriate weights being attributed to different aspects of performance, the validity of assessment information can be threatened (Crooks, Kane, and Cohen 1996). These findings strongly suggest that participation in moderation processes will not, on its own, necessarily enable teachers to learn or to develop suitably valid aggregation approaches. There is thus a compelling case for prioritising further research into how to promote and support the learning necessary for successful social moderation.

With respect to judgement making, the moderation of student assessments is also expected to limit the effects of contextual information (Cooksey, Freebody, and Wyatt-Smith 2007; Harlen 1994b), including teacher biases (Harlen 2005). Teacher biases affect judgement making when assessors draw on knowledge of student abilities or characteristics other than those that are evident in the assessment information (Crooks, Kane, and Cohen 1996; Harlen 2005). In her summary of the findings of a systematic review of research on the reliability and validity of teacher assessment used for summative purposes, Harlen (2005) concludes that student characteristics such as gender, special education needs and behaviour can be sources of bias. Where teachers are evaluating a specific skill, perceptions of overall academic achievement and verbal ability can also influence judgements (Harlen 2005). Likewise, Wyatt-Smith and Castleton (2005) found that teacher judgements of writing can be affected by knowledge of a student’s socioeconomic status or personality as well as by teacher perceptions about the effect that receiving a poor grade might have on a student’s motivation. The difficulty that teachers experience when judging the work of unknown students, about whom it is not possible to draw on additional knowledge, is also documented (Wyatt-Smith and Castleton 2005).

Although teacher bias can be either positive or negative, positive bias presents a greater threat in high-stakes assessment environments where teacher performance might be measured in terms of student outcomes (Crooks, Kane, and Cohen 1996). In New Zealand, the introduction of National Standards has undoubtedly raised the assessment stakes for schools. Changes to reporting requirements that were precipitated by this new assessment system give the impression that it is now possible to compare schools on the basis of student performance. In the run up to the 2011 New Zealand general election, these changes also enabled the National Party to campaign on the grounds that it would shift its school resourcing model to incentivise performance (New Zealand National Party 2011). The National Party has since
been re-elected. It is within this context that National Standards moderation processes are expected to enable teachers to develop the skills associated with not only checking for but also eliminating bias (Ministry of Education 2010c). The underlying assumption here is that because social moderation processes provide teachers with a forum to articulate the conceptions and assumptions on which their judgments are based (Harlen 1994b), these processes must also enable teachers to focus on the qualities and characteristics of student assessment evidence, thereby facilitating the identification and elimination of teacher bias.

Although it is apparent that moderation meetings can provide opportunities to articulate the beliefs and assumptions that inform their approaches to judgement making (Klenowski and Adie 2009; Wyatt-Smith, Klenowski, and Gunn 2010), there is limited evidence that these opportunities actually help teacher moderators learn to identify and address those practices that might be associated with biased judgements. For example, although recent research (Connolly, Klenowski, and Wyatt-Smith 2011) links participation in moderation with the elimination of teacher bias, these findings are based on participant perceptions. Indeed, it has been shown elsewhere that teachers require targeted training to enable them to identify sources of potential bias (Harlen 2005). Similarly, Timperley et al. (2007) note that challenging prevailing discourses, a feature of effective professional learning that is closely linked with the elimination of teacher bias, often requires expert assistance. On this basis, it is contended that where moderation processes are expected to reduce the incidence and effect of teacher bias, teacher moderators will require access to specialist training.

Conclusion

Highly competent assessors bring a set of intellectual and experiential resources to the evaluative process (Sadler 1998). These include deep knowledge of the relevant standards, skills in designing assessments and eliciting responses from students, and evaluative expertise garnered from making comparable judgements in the past. The introduction of National Standards has dramatically altered New Zealand’s assessment landscape. Because of this, the intellectual and experiential resources associated with assessment competency have been redefined. Within this new context, ensuring the efficacy of social moderation processes will necessarily require teacher moderators to acquire a complex range of knowledge and skills.

Although the professional learning opportunities associated with participation in moderation are often promoted (Gardner et al. 2010; Gilmore 2008; Gipps 1994; Harlen 1994a, 2005; Hipkins 2010a; Hipkins and Robertson 2011; Klenowski and Wyatt-Smith 2010a; Linn 1993; Malone, De Lucchi, and Long 2004; Maxwell 2002a, 2010; Radnor and Shaw 1995; Van Krayenhoff et al. 1999), it is evident that simply engaging in these processes does not necessarily provide teachers with the requisite learning opportunities. In addition to obliging teacher moderators to learn new content and skills, effective social moderation requires participants to think about their existing practice in new ways (Connolly, Klenowski, and Wyatt-Smith 2011). The literature on professional learning suggests that any group of professionals would be unlikely to achieve such levels of new learning without the support of someone with the appropriate expertise (Timperley et al. 2007).

Although teacher moderators have identified the need for more extensive training and support (Estyn 2010; Education Review Office 2010), this is not always...
provided. Referring to the 2010 evaluation of arrangements to assure the consistency of teacher assessment in Wales, Daughtery (2011, 7) notes the lack of: ‘reference to the nature or extent of any provision for the training of teachers to prepare them for the central role they have in the system’. It would appear that a similar approach has been adopted with respect to New Zealand’s National Standards. Although the Ministry of Education has provided some professional development to support the introduction of these standards (Ministry of Education 2010c, 2010h), very little of this has been linked with moderation. Indeed, recent research examining the National Standards has identified that schools are struggling to access professional development to help with moderation (Education Review Office 2010), with many school principals of the opinion that moderation processes have been inadequately supported (Thomas and Ward 2011). This finding, which is consistent with schools’ experiences of accessing other professional development in assessment (Absolum et al. 2009), further illustrates the extent to which the demands associated with moderation have been underestimated. It is evident that capacity must be built and supported at all levels to ensure the success of National Standards moderation.

This article is not intended to present social moderation as an untenable process. Neither is its aim to question the role of teacher assessment or the capacity of teachers to produce dependable assessment information. Instead, using the National Standards context as an example, it has sought to highlight the extent and complexity of the professional learning that effective social moderation demands. By drawing attention to the challenges inherent in these processes and questioning assumptions that social moderation will automatically lead to the production of dependable assessment information, its goal has been to lend weight to recent calls for these processes to be better understood and more adequately resourced (Connolly, Klenowski, and Wyatt-Smith 2011; Daughtery 2011; Maxwell 2010). Because developing the skills required to interpret and use assessment information better equips teachers to engage with and respond to student learning needs, professional learning opportunities that include the promotion of assessment capability have been linked to improved student outcomes (Timperley et al. 2007). Effective social moderation thus has the potential to improve both teacher assessment capability and student outcomes. But this is not likely to be achieved unless professional learning is reconfigured as a goal rather than an inevitable by-product of social moderation.

Acknowledgements
The author wishes to acknowledge and thank Alison Gilmore, Lisa F. Smith, Alexandra Gunn and the two anonymous reviewers for their feedback on earlier drafts of this article.

Notes on contributor
Esther Small is a research fellow and doctoral candidate at the University of Otago, New Zealand. She has worked as a primary school teacher and is currently involved in New Zealand’s National Monitoring Study of Student Achievement. Her research interests include social moderation, teacher assessment capability and teacher professional learning.
References


