

Moderation and Teacher Learning

What can research tell us about their interrelationships?

Rosemary Hipkins and Sally Robertson
New Zealand Council for Educational Research



NEW ZEALAND COUNCIL FOR EDUCATIONAL RESEARCH
TE RŪNANGA O AOTEAROA MŌ TE RANGAHAU I TE MĀTAURANGA
WELLINGTON

2011

New Zealand Council for Educational Research
P O Box 3237
Wellington
New Zealand

© NZCER, 2011

Table of Contents

1. About this report	1
Notes about the literature search	1
Key studies on which this synthesis has drawn	2
2. The introduction of National Standards as an imperative for increased moderation activity	5
Making meaning of the standards	6
The need for moderation support	6
Moderation conversations as professional learning	7
3. Some complexities of the moderation process	9
Types of social moderation	9
Complex and active individual and collective decision making	10
<i>Concrete referents</i>	10
<i>Social knowledge as a referent</i>	13
<i>Moderation: A complex and social process</i>	16
4. Moderation as a process of learning together	17
Meaning (learning as experience)	18
Practice (learning as doing)	20
Community (learning as belonging)	21
Identity (learning as becoming)	22
Making sense of the complexity	24
5. Outcomes of moderation	27
Greater understanding of assessment/moderation	27
Increased content/pedagogical knowledge	28
6. Drawing the threads together	31
References	33

Tables

Table 1	Moderation meeting observation checklist	25
---------	--	----

Figures

Figure 1	Components of a social theory of learning: An initial inventory (Wenger, 1998, p. 5)	18
----------	--	----

1. About this report

This report is a synthesis that explores *opportunities for professional learning* that might occur when teachers work together to moderate their judgements about students' work. Since such conversations are inherently social, the synthesis employs the lens of sociocultural learning theory. It draws together key themes from the small number of research projects we found that have documented interactions between teachers as they work together in this way. This introductory section briefly outlines the search strategies we used, and provides overview details of the studies on which we have primarily drawn.

We undertook this review and synthesis to inform a proposed research project of our own. Section 2 outlines why the recent introduction of National Standards has made moderation of students' work an important focus for the professional work of *all* of New Zealand's primary school teachers. It scopes the challenges that teachers are likely to face as they work together to determine whether and how their students' work does or does not meet a specified standard.

Section 3 outlines the complexities that have played out in moderation interactions between teachers elsewhere—complexities that we might thus anticipate New Zealand's primary school teachers are now facing. Section 4 then places these challenges and interactions in a sociocultural learning frame. The synthesis across both chapters has been used to construct an observation checklist for our proposed observation and analysis of moderation conversations in a small number of local primary schools.

Sections 5 and 6 look forward to the potential professional learning benefits of moderation. The central underpinning assumption of the National Standards policy is that the imperative to assess and report against these standards will provide the impetus to lift achievement for students who currently struggle to make the necessary progress in English and mathematics to provide a strong foundation for all future learning. We suggest that there is something of a “black box” between the imperative—laudable as it is—and the impacts that we might anticipate the standards will have. Our overall aim is to inform the implementation of moderation processes in such a manner that our primary teachers will experience rich professional learning opportunities that do support them as they work with their students to lift overall achievement levels.

Notes about the literature search

In an attempt to find research projects that could inform the agenda outlined above we searched using the following combination of concepts: Teacher or Teachers and Moderation or Moderation (Assessment) or Grading or Marking or Students-rating and Professional development or

Learning or Learn or Decision making or Collaborative decision making or Participative decision making.

We searched the following databases:

- Genie—NZCER Library catalogue
- ACER's Ebsco Electronic journals database
- ERIC (Education Resources Information Centre)
- Ebsco Education Research Complete
- Gale Academic Onefile
- Gale Education Reference Complete
- Te Puna National Bibliographic database
- A+Education (Australian Educational Index)
- BEI (British Education Index)
- Google Scholar
- IngentaConnect
- TKI (Te Kete Ipurangi)
- New Zealand Educational Theses database
- ERO website (Education Review Office)

It was not difficult to find material that focused on the provision of professional development for moderation (or grading or rating students—the terminology used differed with each database). However, the proviso that the research should inform us about teachers' actual learning and/or decision making during moderation yielded a much smaller field of relevant papers. Those that most usefully informed our synthesis are briefly described next.

Key studies on which this synthesis has drawn

We found only a small number of empirical studies that discussed the value of social moderation for rich professional learning. Three of these studies (Klenowski & Adie, 2009; Klenowski & Wyatt-Smith, 2010; Wyatt-Smith, Klenowski, & Gunn, 2010) were set in Queensland, Australia, where the New Basics programme provided the impetus for large-scale moderation activities. These studies all appeared to explore the work of the same core group of 15 focus teachers: 12 female primary school teachers and three male secondary school teachers. These teachers were representatives from 15 schools, 11 of which were state schools and four of which were independent. The schools met in clusters and looked at moderation in the English, science and maths areas. The focus teachers were interviewed before and after moderation meetings and meetings were also observed. One paper (Klenowski & Wyatt-Smith, 2010) makes reference to a further 11 schools and 150 teachers (which included 35 focus teachers) who between them moderated the work of students in Years 4, 6 and 9.

Nixon and McClay's (2007) study was much more small scale. It looked at collaborative writing assessment amongst a group of three elementary school colleagues in Canada. The teachers met

every week during their professional development time over three and a half months to discuss their students' written work. The study reports on observations of these meetings and on three interviews with one teacher, who was the focus of the study.

Reid (2007) studied writing moderation meetings in Scotland. Both primary and secondary teachers were part of this action research project. Six teachers were each interviewed twice. The team was involved in a series of planning meetings and a one-day moderation meeting. The moderation process was later extended to other teachers within the cluster group.

We found a New Zealand-based study by Limbrick and Knight (2005). They looked at writing moderation in six primary schools. The 29 teachers involved analysed the writing from the Years 1–8 students within their school first, and then took part in moderation meetings across the cluster of schools. Focus groups for teachers to reflect on the process were transcribed and the teachers were also surveyed about their attitudes towards writing.

There were also a few studies that looked at moderation at the tertiary level. Hunter and Docherty (2011) looked at the moderation process of five tutors from the University of Technology in Sydney. The tutors graded work from a small sample of students in a macroeconomics course. The moderation process involved an initial round of individual grading followed by a discussion of results, followed by a second round of grading and discussion. This study focused on looking at different categories of beliefs held by the tutors.

Kuzich, Groves, O'Hare and Pelliccione's (2010) study investigated moderation of an online unit in the Bachelor of Education at Curtin University in Australia. Four tutors participated in the moderation process. They all had primary school teaching experience but did not have prior experience teaching at the tertiary level. The tutors met in person five times for three hours. The tutors were also interviewed once each and were surveyed. The meetings did not appear to be observed, so the data drawn on in this study came from the surveys and interviews.

2. The introduction of National Standards as an imperative for increased moderation activity

In 2010, National Standards were introduced for students in Years 1–8 in New Zealand schools. The standards specify expected progress in reading, writing and mathematics (Ministry of Education, 2009a, 2009b). This initiative was central to the incoming government’s education policy following the 2008 elections. Parents were promised there would be regular assessment of primary and intermediate students against the new standards, and that reports to them would be in plain English, so that they could clearly see if their child was making expected progress or falling behind. When the latter was the case, targeted funding would help schools provide more support. The overall aim was to improve overall levels of achievement in the foundational areas of literacy and numeracy, and to have all students securely on a learning trajectory that could see them achieve a Level 2 NCEA¹ award in their secondary school years. Following the change of government at the end of 2008, the benchmarks for the standards were developed very rapidly in 2009. They were devised by small groups of curriculum experts in literacy and mathematics, building on previous work that had already established progressions for achievement in these areas.

The progress of Years 1–3 students is to be determined at the end of each 12-month period they have been at school (reflecting the fact that in New Zealand students start school on or near their individual birthday). Progress of students in Years 4–8 is determined at the end of the school year, but teachers are also expected to identify students whose work *during* the year suggests they may not reach the benchmark, and to give them the additional support they need. Most schools had already been using standardised assessment tools before this initiative was introduced. They can continue to use these as evidence of achievement against the new benchmarks but they are also expected to draw on a range of evidence from students’ work in class. Thus schools and teachers face the challenge of using all the available information to make an overall teacher judgement (OTJ) of individual student performance in terms of the standards. In this way the introduction of the standards requires primary teachers and schools to investigate the *meaning* of the work that students generate. While teachers have always made such judgements informally, moderation as an organised process requires making collaborative decisions to reach consensus agreements and hence has become an important professional responsibility for all New Zealand’s primary school teachers.

¹ National Certificate of Education Achievement.

Making meaning of the standards

The standards are very broad descriptors of expected achievement. As such they are going to require a lot of unpacking. Learning from the introduction of New Zealand's standards-based school exit qualification (NCEA) we know that the standard can never reside in words on a page, no matter how carefully thought out and even if accompanied by some examples. A standard is a complex collective of:

- the words used to describe the scope of the standard
- a wide range of examples of tasks that could generate evidence of achievement in relation to this standard
- student work that illustrates the full range we can expect for each task
- an accumulating body of judgements made across this range of work, with particular attention to examples at the boundaries between standards' gradations
- an accumulating individual and collective awareness of all of these aspects within the profession.

Consistency in judgements made using NCEA achievement standards took several years of assessment cycles to consolidate. During these early years the body of available assessment tasks grew. Examiners' reports were produced for externally assessed standards. These reports and feedback from pre- and post-assessment moderation of internally assessed standards helped teams of teachers clarify the intent and expected achievement levels of tasks they had created or adapted from the examples provided. Subject associations and local network clusters contributed to debates and eventually national moderation leaders were appointed for every subject area.² All this moderation activity still requires intensive ongoing maintenance and fine tuning (Hipkins, 2010a). Primary teachers face a time of similarly intense professional learning, for which we can anticipate an immediate need for resources and support.

The need for moderation support

Findings from the 2010 NZCER National Survey of Primary and Intermediate Schools point to an immediate need for moderation support for schools as their practice in relation to reporting against National Standards evolves over time. The principals' responses showed that moderation of OTJs was not yet happening in some schools, even when they were already reporting to parents against the National Standards:

Around 60 percent of the principals whose schools had started work on implementing National Standards said their schools had used OTJs in their mid-year reports to parents. Moderation of writing within a year level occurred in 69 percent of the schools using OTJs for writing in the mid-year reports, and moderation across year levels, in 53 percent of these schools. Around half of schools that used OTJs in reading and mathematics in their mid-year

² Initially teams of teacher moderators did this work in their out-of-class time.

reports to parents also moderated these judgements at the same year level, and around 40 percent, across year levels. This suggests that the OTJs used so far are a work in progress—that OTJs produced within a school may change as the school uses and develops its moderation processes, so that these first OTJs may not be comparable with those made later this year or next year, and so on. (Wylie & Hodgen, 2010, p. 24)

The teachers' responses to the survey suggested only half of them had current experience of moderation in relation to the National Standards. Of this group, around a third had worked with one other teacher, mostly focusing on borderline cases or those where they were unsure. Just over a third of the teachers thought there was a high level of consistency in the OTJs being made across the school and, not surprisingly, they were more likely to be in schools where moderation was already taking place. However, some teachers did not think their school was making consistent OTJs, notwithstanding some moderation activity. This suggests that the quality of in-school moderation is variable. Three-quarters of the teachers and 81 percent of the principals believed that the Ministry of Education (MOE) should support schools in their area to moderate work together to arrive at consistent OTJs (Wylie & Hodgen, 2010, pp. 21–22). This is clearly an area where schools are looking for greater support.

Moderation conversations as professional learning

Hipkins (2010b) contrasts two uses for moderation. Accountability is important when standards-based assessment decisions are high stakes; for example, when the results are used as the basis for awarding qualifications. In these cases, moderation is used to check for consistency in the judgements made and hence to ensure fairness in the overall system. However, moderation can also be seen as an opportunity for rich professional conversations. This review asks how moderation of assessments might present opportunities for professional learning, as opposed to being carried out only for accountability purposes.

What is envisaged is *social* moderation, which has been described as a process that involves teachers discussing and negotiating judgements made about students' work in order to reach a consensus and common understanding of student work standards (Gipps, 1994). While such social moderation is important in ensuring that teachers make reliable and valid judgements in line with those of other teachers and the standards provided (James & Conner, 1992; New Zealand Qualifications Authority, 2007; Wilson, 2004, as cited in Klenowski & Adie, 2009), it can also open up opportunities for professional learning that aims to lift student achievement.

Social moderation takes place in a professional learning community where teachers come together to share ideas and support each other. Learning together as a community can support teachers to take part in processes of “deconstructing, reconstructing and co-constructing knowledge and skills” (Stoll & Bolam, 2005, p. 10, as cited in Millwood, 2007). Beyond the immediate focus on making appropriate moderation decisions, there is potential to build new knowledge about how to more effectively teach so that students have an improved chance of achieving the outcomes targeted by the standards. Improving student achievement is an important aim of any professional

learning activity. According to Timperley (2008), if teachers' professional learning activities are to have an impact on student outcomes, then student outcomes should be the focus of that professional learning. This is a given for social moderation as the focus of the discussion is always on students' work. Moderation can lead to improvements in teaching and learning, for example, by increasing teachers' and students' awareness about what good quality work is like (Ministry of Education, 2010). If this awareness in turn leads to conversations about how to improve teaching practices, then this might result in improvements in the achievement of the students.

This section identifies the *potential* for teacher learning via moderation activities. However, because it is an inherently social process, moderation will involve a range of interactions and decision-making dynamics that may or may not lead to new insights that could usefully inform future teaching. If the potential learning opportunities are to be purposefully leveraged, it will be important to understand as much as possible about the complexity of the decision-making processes that take place. To this end the next section outlines some different types of social moderation and discusses a range of referents that teachers might draw on as moderation decisions are made.

3. Some complexities of the moderation process

Social moderation to reach agreement about student work is not a straightforward process. Since different types of decision-making processes can be used, this section starts by outlining these. Making decisions about the qualities of specific examples of student work involves the use of a number of different resources. This section discusses the use of these in relation to teacher knowledge and the reasoning processes that might be invoked during the making of moderation decisions.

Types of social moderation

Moderation can take many forms. It can take place before, during and after assessments. It is also useful if moderation is ongoing (Maxwell, 2002). Moderation can involve teachers from within a school or across a cluster of schools. It can also take place online, which brings with it challenges associated with the lack of social cues available and sometimes use of the unaccustomed technology (Adie, 2008).

Klenowski and Adie (2009) describe three broad types of social moderation:

- The calibration model: a sample of students' work is graded by teachers individually. The teachers then discuss their judgements with the aim of reaching a consensus and common understanding of the standards.
- The conferencing model: students' work is graded by an individual teacher. Samples of work that represent different levels of performance in relation to a standard are collaboratively selected by teachers and discussed. Again, the aim is to reach consensus and common understanding.
- The expert model: teachers mark all of their work and then submit it to an expert. Teachers receive feedback on whether the standards have been interpreted and applied in the way in which they were intended (Queensland Studies Authority, 2007, as cited in Klenowski & Adie, 2009).

Some moderation processes are more formal than others: for example, some are managed by experienced facilitators while others consist of a group of teachers meeting informally with one another. Moderation discussions can take place during different assessment phases, sometimes just involving the checking and confirming of post-assessment decisions, and at other times involving the development of the assessment tasks and the marking criteria.

In practice, hybrids of different moderation models may be employed, or over time the emphasis could shift from one model to another.

Complex and active individual and collective decision making

It is clear when looking at the literature that social moderation is a very complex process, with a number of steps involved. To begin with, each teacher must arrive at the personal judgements they bring to any subsequent moderation conversation with their peers. Then, working together, they need to compare individual judgements to arrive at shared understandings, with informed agreement between them as the ultimate aim of moderation.

Sadler (1998, p.80, as cited in Wyatt-Smith et al., 2010) described three sequential components to the judgements a teacher makes about students' work. These are: the teacher paying attention to the learner's work; assessing the work against some sort of standard; and making a response (for example, grading the work). At each decision point different teachers might draw on different types of resources when making the relevant judgement. Depending on the details to which they attend, and the resources to which they pay relatively greater or lesser attention, different teachers could come to different conclusions about the same or similar pieces of student work.

A number of studies have looked at how teachers make use of available resources as they make their decisions. The complex mix of resources drawn on and decisions made are now discussed. First we outline how resources relevant to the judgement to be made are actually used, and following that we investigate social knowledge as another type of referent that is potentially available for teacher decision making.

Concrete referents

Teachers can access a number of concrete referents to assist them in making moderation decisions and they use and prioritise these in different ways. Concrete referents include:

- the standards themselves
- any guidelines or other guidance material provided
- the student work to be moderated.

The assessment tasks are also potential referents, and they may well be subject to pre-assessment moderation. However, the role that actual tasks could play in moderation is beyond the scope of this review, which takes as its immediate focus the meaning that teachers infer from work that students have already produced.

Standards as referents

Standards describe the requirements and expectations for students at particular levels. They can help teachers to focus on what is being assessed. Maxwell (2008, p. 2) identifies five different types of standards:

- standards as moral or ethical imperatives (what someone should do)
- standards as legal or regulatory requirements (what someone must do)
- standards as target benchmarks (expected practice or performance)
- standards as arbiters of quality (relative success or merit)
- standards as milestones (progressive or developmental targets).

All standards are open to interpretations of meaning, intent and the making of value judgements, so all of these types of standards might require moderation conversations at some point in decision making. However, the focus and content of the conversation in each case is likely to differ. The National Standards are target benchmarks of expected performance by a specified age or year level. The focus of conversations is likely to be on whether or not a specific piece of work does show evidence of having reached the benchmark specified.

As mentioned in Section 2, written statements on their own are not enough to define standards or to produce consistent teacher judgements. It is not possible to capture the richness of a construct through a description alone. Instead, standards get their meaning from being used over time. “Shared professional knowledge about what constitutes the standard builds slowly but continuously as it is enacted” (Hipkins, 2010b, p. 19). Therefore, the moderation process makes an important contribution to the meaning that the standards come to gain.

In part, shared meanings evolve as teachers discuss the meaning of the words used to specify a standard. It is important that key terms are defined and interpretations agreed as this enables teachers to work together from the basis of a shared understanding and discourse. Such common *understandings* can help to produce more reliable teacher judgements and it seems likely that teachers will be aware of this (see, for example, Wyatt-Smith et al., 2010).

Standards can vary in how general or detailed they are. Although a common understanding of the standards is important, highly prescriptive standards have their drawbacks. If the criteria are too specific, “achievement can come to be seen as little more than criteria compliance in pursuit of grades” (Torrance, 2007, p. 286). Such compliance behaviour has been widely critiqued as not supportive of improvements in students’ learning. Less prescriptive standards, such as the National Standards recently introduced for New Zealand schools, allow space for meaning to be developed over time, and they have a better chance of contributing to changes in teaching and learning that do ultimately lift achievement levels.

Once a shared understanding of the described standard has been reached, questions of how the standard should be *applied* must also be debated. Standards can be interpreted differently by different teachers and by the same teacher in different situations (Sadler, 2008, as cited in Adie, 2008). For example, a standard might be seen as describing either a minimum requirement or a typical performance (Klenowski & Adie, 2009). Teachers may not know which interpretation is pertinent without the relevant guidance. This suggests that information on how to apply the standard should be provided to teachers, with the further implication that discussions amongst the teachers may be needed to clarify the intended meaning of such advice.

Guidelines and other professional advice as referents

Guidance, including teacher guidelines and other official sources of professional advice,³ may be provided to support a set of standards. Such guidance could provide direction for making judgements about the individual components of the assessment (the micro level) as well as information on making an overall judgement (the macro level). As with the standards themselves, teachers need to have a common understanding of the terms used in guidance materials. When they do not share common interpretations teachers may be inconsistent in their use of terms and hence apply advice in different ways.

Exemplars of student work may or may not be included in guidance materials. These are annotated samples of students' work with an explanation about how the standard was applied in each case. However, these can only ever be *illustrative* of a standard because there will be many possible ways in which students can demonstrate that the standard has been achieved.

Teachers use and prioritise the concrete referents available to them in different ways. For example, during social moderation meetings, teachers may vary in the priority they give to a specific piece of guidance. Some teachers in the New Basics study saw the guide that had been developed as their primary reference, while others did not prioritise the guide at all (Wyatt-Smith et al., 2010). When the student work did not align easily with the standard or was on the boundary between two levels of a standard, teachers were more likely to refer to the guide. Teachers can also vary in how much attention they pay to the exemplars. In the New Basics study some teachers found that annotated samples were a good way to check their marking process, particularly when the overall grade was on the border of two levels. Some teachers relied heavily on the exemplars while others reported that they did not find them helpful (Wyatt-Smith et al., 2010).

While the guide and exemplars are designed to help teachers to make moderation decisions, sometimes they, along with the standards, can compete for teachers' attention rather than complementing each other as intended. When teachers in the New Basics study had initially recorded different grades they used a variety of strategies to resolve their differences during the subsequent moderation meeting (Klenowski & Adie, 2009). Some looked at the description of the standards, some teachers used both the guide and the standards and some teachers capitulated to group pressure (usually when the disagreement was over an element that did not affect the overall grade). The teachers' study perceived that all the different materials complicated the judgement process and they lacked "a coherent frame for guiding their decision-making" (Wyatt-Smith et al., 2010, p. 65) even though the different pieces of guidance should have been able to be used to complement each other. This suggests that information provided to teachers should include

³ New Zealand's Ministry of Education has published a number of commentaries to support standards implementation in the *Gazette/Tukutuku Korero*—a weekly magazine to which all schools have ready access. The online portal www.tki.org.nz also provides officially sanctioned resources and advice, including examples of good practice from leading schools.

discussion of how the provided materials are related and might most productively be used in conjunction with each other.

Other aspects of student work as a referent

Both individually and collectively, teachers may need to:

- make moderation decisions about individual pieces of student work
- make an overall judgement in relation to multiple pieces of work
- decide the relative weightings to give to different aspects of the same piece of work.

We found few studies that discussed in detail what teachers choose to focus on and notice about the work itself. However, some studies did discuss the different ways teachers move from making decisions about the separate parts of the work (the micro level) to making overall level judgements (the macro level).

Teachers can be given different instructions about how to make overall judgements. Sometimes these instructions are not explicitly given and teachers use different strategies to switch between looking at parts of the students' work and the whole. Klenowski and Wyatt-Smith (2010) document that some teachers totalled subscores for parts of the assessment task, some prioritised looking at the exemplars and some used a different group of teachers to assess the overall score from the group who assessed each component. Klenowski and Adie (2009) similarly reported that there were different ways of coming to an overall grade (e.g., averaging the grades for the elements, judging the task as a whole, placing different emphasis on different elements).

Ultimately, teachers need to decide which level of a standard is the best fit for the student work and sometimes the work can display features of more than one level (and in some cases contribute evidence of achievement relevant to more than one standard). The decisions to be made are not straightforward, no matter how well teachers might understand the provided referents. Wyatt-Smith et al. (2010) noted that when the decision was not clear-cut, teachers used compensations or trade-offs to make a final decision, as discussed next. The reasons they might give for making such value judgements point to the contribution that social knowledge also makes to teachers' decision making.

Social knowledge as a referent

Social knowledge (e.g., beliefs and past experiences) comprises an important part of teachers' working knowledge. This working knowledge in turn influences the choice of evidence and then how the selected evidence is interpreted (e.g., Coburn & Talbert, 2006, as cited in Coburn, Toure, & Yamashita, 2009). As Gipps explained:

We are social beings who construe the world according to our values and perceptions: thus our biographies are central to what we see and how we interpret it. Similarly in assessment, performance is not 'objective'; rather it is construed according to the perspectives and values of the assessor. (1999, p. 370, as cited in Scarino, 2005).

If we accept Gipps' proposition, it follows that teacher judgement is necessarily subjective and that moderation decisions made by teachers need to be *defensible* because it is not possible for them to be unproblematically right or wrong in any singular objective framing. It follows that recognising and encouraging discussion about social knowledge is important for decision making during moderation (Adie, 2008). Such discussion can help teachers to recognise and challenge their own and others' assumptions as they work towards a consensus of the meaning of work in relation to a standard.

Many different categorisations are used in the literature to describe the different social knowledges on which teachers might draw. We have chosen to use just three overarching categories of knowledge and beliefs. These are:

- knowledge and beliefs about assessment
- knowledge and beliefs about students
- knowledge and beliefs about the intended curriculum.

In what follows it will be evident that even these broad categories have a considerable degree of overlap.

Knowledge and beliefs about assessment

Teachers' *prior experiences of making judgements* are likely to inform the methods they think should be used to reach overall scores. Their views about the *consequences* of assessment are likely to inform decisions about the range of levels that should be awarded. For example, if a teacher believes that reporting that a student has not reached expected levels of achievement would be highly likely to discourage this student from making further effort, they may not want to make this judgement—at least in public (Hipkins, 2010c). A related belief is that any effort should earn a student some recognition (e.g., McMillan & Nash, 2000; Wyatt-Smith et al., 2010).

Teachers might also question the *quality* of the assessment task and/or the indicated parameters for work that meets the standard. For example, some teachers in Wyatt-Smith et al.'s (2010) study disagreed with the focus of the standards. Based on their experience, they would have focused on other aspects of the task. Dissent of this sort is likely to impact on how teachers apply (or do not apply) the indicated standard. They might, for example, look for responses that are in line with what they think is a good response (rather than what the standard actually specifies) (e.g., Klenowski & Adie, 2009).

Knowledge and beliefs about students

Teachers are likely to be influenced by their knowledge and beliefs about particular students as well as about their beliefs and expectations of students in general. Some beliefs/knowledge that can influence teachers' moderation decisions at general and/or student-specific levels include:

- whether or not assessment should recognise effort (as well as achievement)

- whether assessment would demotivate some students by identifying and documenting poor levels of achievement relative to their peers and to expected levels of progress
- how conditions under which an assessment was completed might have influenced performance
- beliefs about what an “average” student should be able to achieve
- perceptions of the ability of individual students, or groups of students
- knowledge of specific challenges student(s) might have faced.

In the New Basics study a number of teachers used their knowledge of particular students’ attributes and dispositions to influence their decisions (Wyatt-Smith et al., 2010). However, there was a range of views about how appropriate it was to draw on this knowledge. The teachers in their study saw the benefits of applying standards differently in different cases, based on their knowledge of the students. Klenowski and Adie (2009) also found that there were different opinions about how objective teachers thought they could and should be. On the basis of their research, Klenowski and Wyatt-Smith (2010) suggested a need to clarify the difference between making decisions based on achievement and decisions influenced by students’ attitudes (such as diligence and disposition).

Knowledge and beliefs about the curriculum

Gill and Hoffman (2009) found that teacher beliefs about the following influenced their grading of students’ work:

- pedagogical content knowledge (PCK)
- general pedagogical beliefs
- subject matter beliefs; beliefs about curriculum
- beliefs about the role that textbooks should play in learning.

Note that this study focused on grading (marking) decisions rather than moderation per se. However, the list of influences is a useful pointer to the knowledge that might also be invoked in making moderation decisions. Arguably the common threads that bind these potentially very large belief sets together are beliefs about the nature of learning and what schooling should be for—that is, the *purposes* for learning in general and specific curriculum components in particular. It is likely that such views will be tacitly held rather than transparently available to inform decision making so it is important that their impact on decision making is explored and documented.

Hunter and Docherty (2011) found that moderators were better able to agree on criteria within their discipline area rather than on the criteria that related to generic skills such as communication. This suggests that the former (subject purpose) was at least tacitly understood in the same way by all of them but the latter (additional or new purposes for the learning) was seen as open to being contested.

Notwithstanding the availability of a common guide, the science and mathematics teachers in the New Basics study used different judgement processes from those used by the English teachers (Klenowski & Wyatt-Smith, 2010). These differences related to their beliefs about how desirable/possible it was to be objective in their decision making. The mathematics and science

teachers believed that standards should be detailed and objective. Their preference was to arrive at an overall grade by adding the marks for each element of the assessment task. By contrast, the English teachers were more inclined to look at the work as a whole. They were more likely to be supportive of “fuzzy” standards, and were more likely to believe that standards developed over time and through interpretation.

In a survey of the assessment practices of 513 high school teachers in Canada who taught a range of subjects, Duncan and Noonan (2007) similarly found differences in the decision-making process for teachers of different subjects. Compared to other subject teachers (science, English, social studies) mathematics teachers did not give as much preference to considering noncognitive aspects such as student effort and motivation.

Moderation: A complex and social process

The view that assessment is a technical process rather than one based on interpretation and judgement represents a traditional, positivist paradigm (Davidson, 2004, as cited in Scarino, 2005). In this paradigmatic way of thinking, achieving “objective” decisions is the ideal to which moderation aspires, even if reality suggests it is not easy to achieve. However, it should be clear from the discussion in this section that this paradigm does not serve as an adequate basis for considering moderation decision making when there are so many variables to take into account, and when they can interact with each other in so many different ways. We need a different theoretical frame for thinking about the complexities of social learning and decision making.

Within an interpretivist paradigm, assessment is seen as a social process that is underpinned by cultural beliefs and relationships (Scarino, 2005). In this theoretical frame, explicitness, transparency and training are valued (as they are in a positivist framing) but the role of interpretation is also acknowledged as important (Scarino, 2005). The process is necessarily subjective but it is important not to take this as meaning that “anything goes”. Instead of subjectivity being seen as undermining validity and reliability, it should be framed as being part of the process. Social moderation can help by enabling teachers’ tacit knowledge to be made explicit (e.g., Maxwell, 2002) and to ensure those taking part are mindful of the role of the different referents as they make moderation decisions (Klenowski & Adie, 2009). The next section addresses the challenges of developing processes that can support this type of socially-located professional learning.

4. Moderation as a process of learning together

Section 2 began to build a picture of the complexities of moderation by outlining a range of ways in which referents might be accessed and deployed when making decisions about student work. This section adds to this picture of complexity by exploring the moderation process as a social activity in which a group of people learn together. Within such a framing, new learning possibilities can emerge in the spaces between the individuals who are taking part. Rather than simply pooling all that they know—an additive model for learning together—knowledge that none of the participants might have held before could come into view. This model frames learning conversations as interactive and emergent. As ideas are wrestled with and differences resolved, new insights can become apparent, and hopefully will ultimately become shared if they are seen to be improvements on previous professional knowledge.

Wenger's (1998) social theory of learning provided us with a helpful frame for considering the complexities entailed here. Four components are central to this theory:

1. Meaning—a way of talking about our (changing) ability—individually and collectively—to experience our life and the world as meaningful.
2. Practise—a way of talking about the shared historical and social resources, frameworks, and perspectives that can sustain mutual engagement in action.
3. Community—a way of talking about the social configurations in which our enterprises are defined as worth pursuing and our participation is recognisable as competence.
4. Identity—a way of talking about how learning changes who we are and creates personal histories of becoming in the context of our communities (p. 5).

These four components can be seen in Figure 1 which places learning at the centre of these four interconnected components: meaning (learning as experience); practice (learning as doing); community (learning as belonging); and identity (learning as becoming).

Figure 1 **Components of a social theory of learning: An initial inventory (Wenger, 1998, p. 5)**



Social moderation involves all four components of meaning, practice, community and identity in the context of the development of a “community of practice”. In turn, a community of practice involves:

- mutual engagement of the members (members building collaborative relationships)
- joint enterprise (a shared understanding of the purpose of the community of practice)
- a shared repertoire (communal resources) (Wenger, 1998).

In the context of social moderation, a community of practice includes the teachers involved in the moderation meeting and anyone else who works with them to support their decision making. Teachers gain experience in moderation by participating in this community of moderators (Lave & Wenger, 1991, as cited in Klenowski & Adie, 2009).

The four components of meaning, practice, community and identity are now discussed in relation to the assertion that a group of teachers carrying out moderation together can constitute a community of practice, given appropriate focus, flow and facilitation of their shared conversations and individual reflective learning.

Meaning (learning as experience)

As Section 2 outlined, moderation involves the sharing, negotiation and construction of meanings in order for the resources used to come to decisions; that is, concrete referents that include but are

not limited to the standards, students' work and any guidelines and guidance that have been provided. Teachers bring their own working knowledge to the manner in which they draw on these referents during moderation and it is through the *experience* of sharing and negotiation that shared meanings and understandings can be created. "Meanings are produced which may support, enhance or challenge current understandings and values held by the teachers" (Klenowski & Adie, 2009, p. 20). The possibility of challenge makes it likely that conflict and tension could be a part of the process that teachers experience (Wenger, 1998).

Section 2 also noted that moderation requires the development of a shared language so that teachers do not talk past each other. Teachers need to discuss key terms so that they can agree on how to grade the work and have a common understanding about what the standards mean. This is essential given that the ultimate aim of moderation is to reach a consensus. However, a sociocultural framework adds to the challenges of doing this by pointing to the part that cultural and historical contexts play in the construction of meaning (Murphy, 2000, as cited in Nixon & McClay, 2007). Teachers need opportunities to work together as a "community of interpreters" (William, 1998, cited in Reid, 2007). If they do not have opportunities to work in ways that surface the working knowledge they each bring to the conversation, differences are unlikely to be resolved and they are unlikely to learn why these differences existed to begin with. This challenge points to the necessity of making space for individual reflection as well as for building group understandings. Teachers need to learn more about themselves as well as to learn from each other.

Moving between group and individual reflection suggests that moderation is best structured as an iterative process where shared understandings grow over time. We should not anticipate that "one-off" meetings will suffice. Similarly, the provision of opportunities for reflection at different stages of the assessment/moderation sequence is seen as helpful. Saunders and Davis (1998, as cited in Kuzich et al., 2010, p. 4) describe a sequence involving three steps if moderation processes are to result in greater marker consistency as a result of common understanding. These are:

- initial discussion about the meaning of the criteria
- further discussion to refine understandings after some marking has been undertaken
- moderation discussion after all the marking has been completed.

Reid (2007) found that, initially, teachers were focused on gaining more detailed interpretations of the criterion statements. When the facilitator helped teachers to discuss the aspects such as linguistic principles or the meaning of phrases such as "complexity of thought", discussions became more collaborative. As Reid noted, "when discussion was underpinned by a shared understanding of a principle for progression, joint discussions were more fruitful" (p. 140). This points to the importance of building meaning for deeper matters related to the purposes and qualities of learning rather than for surface-level indicators of successful task completion. In this way, the experience of making meaning during moderation is curriculum work, at least as much as it is assessment work.

Practice (learning as doing)

The most obvious consideration here is practice itself. Opportunities to practise new skills require ongoing experiences of moderation rather than “single event” learning opportunities (Strachan, 1996). A related consideration is that the working environment must feel like a safe place to practise. Sharing and challenging one’s own and others’ beliefs can make people feel vulnerable and anxious so it is important that practice opportunities are conducted in a trusting and supportive environment. However, time expended in working together is necessary for such trust to be established. Paradoxically, if teachers do not feel this time is, or could be, time well spent, the desired productive learning climate might not be readily achieved. Illustrating this dilemma Klenowski and Adie (2009) found that the time required for the moderation process was viewed negatively, even though the teachers recognised that they needed time to practise in order to feel confident about sharing their views. An implication here is that practice needs to demonstrably lead somewhere constructive. Teachers, no less than students, need to feel they are making progress in their learning, and that what they are learning has personal salience for them. This would seem more likely to happen if moderation is focused on improving learning rather than on compliance.

Nixon and McClay (2007) identified four phases to the flow of discussion within one moderation meeting: group gossiping/spectating; reading/rereading; deliberating/reframing; and collaboratively creating. Group gossiping/spectating involved the teachers talking about each student’s identity and about the writing that took place. They talked about their beliefs and values about the student, the task and about the assessment task. The teachers then shifted to using more transactional language as they moved away from group gossiping to group reading/rereading. During this phase the teachers would read the student’s work, adding their reflections on the work and thoughts about the student’s identity and purpose. In the deliberating/reframing phase, the teachers used the scoring guide and focused on making sense of the rubric. The teachers discussed, persuaded and disagreed with one another until a consensus was reached. A number talked about how it was not possible to put aside knowledge of the students during this marking. Reframing of their understandings occurred as they reflected on each other’s assumptions. Such rich descriptions of actual practice point to a purposeful flow and logic to moderation conversations. An implication is that taking “shortcuts” may not lead to the desired outcome of developing a consensus that is deeply informed by insights that emerge during the collaborative conversation. These conversations really do need a meaningful block of time.

Strong facilitation by a leader who is well informed about productive conversation sequences will doubtless help to pace conversations in ways that maximise learning opportunities yet avoid frustration for participants. Skilled facilitators also have a role to play in the construction of a safe learning environment. Healy and Bush (2010) suggest that it is important to have protocol for guiding moderation discussions that enables teachers to share their views equally and so that discussions remain focused. Guided discussions can help create a collegial and safe environment where sharing can take place.

Community (learning as belonging)

Social moderation ideally involves belonging and participating in a community whose members feel safe to support and challenge one another. Their aim as they work together is to reach a consensus about the meaning of student work in relation to the specified standard(s) and this will sometimes require at least some of the community to reshape their current beliefs:

New meanings, new behaviours, new skills, and new beliefs depend significantly on whether teachers are working as isolated individuals or are exchanging ideas, support, and positive feelings about their work. (Fullan, 2001, p. 84, as cited in Barrett, 2008)

This consideration further reinforces the importance of spending time to build a sense of community and belonging if teachers are to experience moderation as a process of ongoing professional learning. However, it cannot be assumed that simply investing the necessary time will be sufficient to ensure that a sense of belonging does emerge. Limbrick and Knight (2005) reported that a number of the New Zealand teachers in their study experienced moderation as not being collegial in nature, with a process that was frustrating and/or tense. Some teachers did not think there was professional trust and respect within the group, and a lack of professional confidence was apparent in some interactions. Some teachers were unhappy with the size of the group. Others commented that there seemed to be different expectations about the process of moderation. Nevertheless, many of the teachers found the process to be a positive one and the opportunity to have collegial discussions to be useful. Studies such as this are an important reminder that different teachers can experience the same conversation quite differently.

One reason for the experience of belonging being different for different individuals could be that whose voice is heard depends on the particular personalities, dynamics and aims of the moderation community. For example, Klenowski and Adie (2009) reported that when there were disagreements, teachers often chose to favour the grade given by the student's own teacher over that given by other teachers. Here the teacher's working knowledge determined whose voice was given priority. At other times, the political and organisational context in which a conversation takes place results in particular interpretations being valued over others. In a study about the use of evidence in decision making, Coburn et al. (2009) found that "evidence" might be used symbolically to support decisions that were actually predetermined by the particular power dynamics of the context. (Note that this was not a moderation study but it did require a group of teachers to work together to make a policy decision.) Observations about group dynamics such as these again point to the need to carefully consider whether or not a skilled facilitator should be a part of the group.

Using skilled facilitators and making increasing use of experienced teachers as facilitators can increase the likelihood that meanings can be challenged and negotiated without inadvertently alienating any members of the group. However, few studies mention this role in much detail. Facilitators can guide the analysis process, challenging assumptions and suggesting new ways of looking at/doing things (Timperley, 2008). For example, Hunter and Docherty (2011) described a moderation meeting run by an instructor who had been involved at all stages of the professional

learning programme related to the student work being discussed. Critical dialogue was encouraged and the facilitator questioned different teachers about their views in order to highlight different values that underpinned their arguments. The facilitator also drew attention to inconsistencies; for example, between markers' grades and their comments. Earl and Katz (2003, as cited in Millwood, 2007), similarly note that "external critical friends" can help teachers by asking questions and facilitating reflection as well as help in the analysis and interpretation of complex data. In their study, the teachers themselves were aware that working with a facilitator made an important contribution to the success of the moderation process. These studies point to the demanding nature of the facilitator role, and to the deep knowledge needed by those who work with teachers to build a sense of community and shared purpose around understanding the meaning of work that students generate.

Another important consideration in the context of "learning as belonging" is determining who can and should belong. As well as deciding whether or not to involve a facilitator, consideration needs to be given to ensuring that members of the group will bring diverse perspectives as grist for the conversation (Hipkins, 2011). According to Price (2005, as cited in Kuzich et al., 2010), both expert and novice markers should be a part of the assessment community because all the teachers need to be socialised within the particular learning community. Reading between the lines, this suggests that both expert and novice markers can contribute diverse perspectives to the moderation conversations and there is potential for both the expert and novice markers to learn from one another. There is also a clear implication that moderation tasks should not be assigned only to the more experienced teachers (as might happen if only efficiency is being sought).

Moderation need not be limited to teachers within the same school. Klenowski and Adie (2009) reported that the teachers in their study appreciated the opportunity to talk to teachers with whom they would not otherwise interact. Students might also become part of the community of learners. For example, Reid's (2007) study initially involved only teachers but the community was later extended to include students in the process.

Beyond what is reported here, we found little documentation of the impact of group membership on the learning dynamics. This is clearly a topic in need of further attention if our aim is to get better at designing moderation conversations that do successfully lead to teacher change and professional growth.

Identity (learning as becoming)

Seeing oneself *as* a learner is important for changing teacher practice (Brookfield, 1995, as cited in Millwood, 2007). Millwood found that learning from the moderation process involved teachers coming to see themselves as learners who were open to new ideas and willing to challenge their current thinking. These observations illustrate how, from a sociocultural perspective, learning involves a "transformation of identity, where identity is understood as evolving forms of competence" (Murphy & Hall, 2008, p. ix). This sense of becoming might entail building stronger

competencies as a moderator and/or becoming more skilful and responsive as a teacher. In this review we have aimed to bring both identities (moderator and teacher) together with an ultimate focus on improving student learning.

How then, can teachers best be supported so that they do come to view themselves initially as apprentice moderators, and increasingly as informed and skilful judges of student work? Clearly, participation in appropriately structured moderation meetings is important. However, coming to view oneself as a moderator can take time. Klenowski and Adie (2009) found that all of the teachers in their study identified themselves as apprentice moderators when the standards were new to them. These researchers suggested that coming to be recognised by their community as experienced moderators does require considerable practice and time. However, practice per se, while necessary, is not likely to be sufficient for moderation competencies to grow in ways that do impact on subsequent teaching. The teachers in Nixon and McClay's (2007) study, who did come to see themselves as successful moderators, engaged in pedagogical self-evaluation and critical reflection as a part of the moderation process. The implication here is that opportunities for reflection are important for building teachers' sense of becoming competent evaluators of student work. Again, the importance of skilled facilitation is highlighted by this implication.

Teachers bring particular worldviews to moderation meetings. For example, their identities as teachers of particular subjects, as primary or secondary teachers and as experienced or beginning teachers have an impact on their moderation practices, but these identities might also be challenged and modified in positive ways as a consequence of their learning experiences during moderation. Illustrating this potential, Reid's (2007) study involved primary and secondary teachers working together, with each group bringing their different experiences of teaching and of moderation. The primary teachers were able to provide information about the contexts in which the student work was produced. They felt they had improved their subject knowledge as a result of their interactions with secondary teachers. The secondary teachers provided subject knowledge to the conversations and felt they had learned about pedagogy from the primary teachers.

Relative levels of teaching experience are likely to impact on interactions and decision making during moderation meetings, and hence on teachers' sense that their professional expertise is growing as a result of moderation activities. Illustrating this, Hunter and Docherty (2011) found that relatively inexperienced teacher assessors were open to negotiation and to developing their knowledge about writing through the application of the writing criteria specified. By contrast, teachers who were more experienced found the criteria too specific, and hence less helpful either for making moderation decisions or for expanding their pedagogical knowledge. Similarly, Klenowski and Adie (2009) found that those teachers with more teaching experience tended to use the standards provided as a guide, whereas those with less experience were more likely to use the standards more exactly.

Making sense of the complexity

Moderation not only involves drawing on a complex mix of resources, but also entails complex social interactions. When moderation meetings occur, the ideas, beliefs and resources that the teachers bring can interact in ways that allow new meanings and understandings to emerge. The research challenge here lies in observing and attempting to better understand what happens in the *spaces between* those involved in the moderation, not just in exploring what individual teachers bring to the moderation process.

In order to be able to study the complexity of interactions within the moderation process, we have used the insights from the studies reported above to develop a tentative checklist of aspects that would be useful to pay attention to when observing moderation meetings (see Table 1). We still have much to learn in this area and therefore this checklist is very much a work in progress.

Table 1 **Moderation meeting observation checklist**

Concept	Key questions
Meeting structure	<p>How much time is given to the meetings?</p> <p>Where do they take place? (conversation conditions)</p> <p>How is the group led? Is there a facilitator? If so, what is their role?</p> <p>Is there any protocol guiding the moderation discussions?</p>
Group identity	<p>Who is included in the group (e.g., experienced/inexperienced moderators, same/different school, subjects taught, primary and/or secondary teachers)?</p> <p>Are students involved in any way?</p>
Trust	<p>Does anyone appear to be anxious?</p> <p>How does the group establish trust? (Does the facilitator create a safe learning environment?)</p> <p>(How) do group members support one another?</p>
Active engagement	<p>Who is contributing? What sorts of contributions do they make? (observations, questions, comments)</p> <p>What is the focus of each contribution? (students' work, standard, guidelines, a specific student)</p> <p>Do discussions remain focused? Is there any process or support for this to happen?</p> <p>How invested in the process are the group members?</p>
Shared meanings	<p>(How) are meanings shared?</p> <p>(How) does the group arrive at a consensus?</p> <p>Whose views do others challenge and whose do they support? On what grounds do they do so?</p>
Power dynamics	<p>Whose voice is heard/not heard?</p> <p>How are suggestions treated?</p> <p>Whose views are prioritised when reaching a decision?</p>
Deep learning	<p>Are individual members and the group given time for critical reflection? What occurs during this reflection time?</p> <p>Is discussion focused on surface-level features or the deeper purposes and qualities of learning?</p> <p>Do group members challenge their own and others' beliefs?</p> <p>Is there an emphasis on learning (as opposed to compliance)?</p>

5. Outcomes of moderation

In this review, we have searched for insights regarding how best to ensure that moderation processes can inform teaching and learning with the ultimate goal of improving learning outcomes for students. Improved outcomes are more likely to follow when moderation provides a springboard for conversations about the content of learning, teaching strategies to use and how best to address the specific needs of students (Healy & Bush, 2010). In order for moderation to have an ongoing positive impact, it is not just what happens in the moderation meetings that matters, but also the organisational support that is put in place to help teachers to act on new insights during subsequent teaching and learning.

Our search for studies that could inform our own proposed work confirmed that there is a paucity of research that focuses on the *learning* of teachers through social moderation, let alone on the subsequent translation of new professional insights into teaching. Despite a number of studies conducting post-moderation interviews with teachers, few studies appear to have focused on the impact of moderation on the classroom. Klenowski and Wyatt-Smith (2010) noted some teachers in their study did not see how to apply the standards to teaching and learning activities in the classroom, despite having recognised the impact of the standards on helping them to focus on important areas of assessment.

Why is it that moderation does not necessarily lead to improvements in teaching practices or improved student achievement? Research in related areas suggests that transferring what has been learned through the moderation process to the classroom is unlikely to be a simple process. For example, Parr and Timperley (2009) have raised questions about the preparedness of schools and teachers to engage with the meaning of assessment results. One potential issue, as we have seen in earlier sections, is that there may be different opinions about the form evidence should take (Timperley & Parr, 2005, as cited in Parr & Timperley, 2009). Another issue is that teachers may not have the knowledge and skills required to work with student achievement data. Parr and Timperley (2009) conclude that, in order to have an impact on student achievement, teachers also need pedagogical content knowledge so that they can teach to meet students' needs. All of these issues raise interesting possibilities for further exploration of the potential for achieving outcomes from moderation that could result in strengthening classroom learning.

Greater understanding of assessment/moderation

One of the potential outcomes of well-facilitated moderation conversations is that teachers gain a greater understanding of *both* assessment and moderation processes, and feel more confident of

their abilities to judge the meaning of work generated during class as a result. For example, the tutors in a study conducted by Kuzich et al. (2010) reported that after taking part in moderation processes they had a better understanding of assessment and of what different student responses looked like at different levels. Reid (2007) similarly found that the teachers involved in moderation were more confident in making assessment judgements, which in turn led to them being confident to talk to their students about what strong evidence of achievement could look like, showing them exemplars and making links between assessment criteria and their teaching goals.

Some teachers in the New Basics study (Klenowski & Wyatt-Smith, 2010) thought that the standards they were using would help them to support students to focus their work on the things that were being assessed and to understand their strengths and weaknesses. However, some teachers did not appear to see these opportunities. The researchers concluded that there is a need to build teachers' capacity in making links between assessment and *learning theory*, specifically the role played by self-assessment in strengthening learning and lifting achievement. An implication here is that a sociocultural framing of moderation processes, as outlined in Section 3, could productively be an overt focus of professional learning conversations. This could support teachers to enrich their understandings of their own and their students' learning processes as a direct result of taking part in well-facilitated moderation conversations.

There is evidence that conversations that involve students in making informed assessment decisions can be a very effective means for teachers to lift achievement levels (Absolum, Flockton, Hattie, Hipkins, & Reid, 2009) so strengthening teachers' confidence and knowledge to regularly hold such conversations with their students would appear to be an important outcome of moderation to be intentionally fostered.

Increased content/pedagogical knowledge

Teacher development activities have the aim of supporting teachers to change or take on new knowledge, beliefs and skills in order to make changes to their teaching practices. "Professional development involves not only the use of teaching activities in the classroom but also the development of the beliefs and knowledge underlying the practice" (Fazio & Gallagher, 2009, p. 4). As the earlier sections have outlined, effective leadership of moderation meetings can provide the space for teachers to reflect on their own current content knowledge as well as learn from the knowledge and reasoning of others. A positive outcome here would be an increase in each teacher's awareness of areas in which they may need to develop their personal professional knowledge and skills. If they use their increased knowledge and awareness to further their pedagogical and content knowledge this should in turn contribute to improved teaching in the classroom, and therefore has the potential to help lift student achievement.

Illustrating this change potential in action, the teachers in Limbrick and Knight's (2005) research reported that moderation helped them to become more aware of their own knowledge about

writing and the writing process, including understanding of how to recognise and interpret deep features of writing (e.g., impact, voice and audience). Many reported that they were now aware of gaps in their knowledge but, overall, they had an increased confidence in their ability to teach writing. The teachers identified the modelling of the facilitator and opportunities to share pedagogical knowledge with other teachers as factors that allowed them to learn together, to reflect on their own practice and to take their new knowledge and confidence to the classroom. The researchers noted reports of instances where key terms and features were subsequently used by the teachers when working with students. These practices were helping them make their teaching of writing more explicit for the students.

6. Drawing the threads together

At the start of the review we noted that New Zealand's teachers of students in Years 1–8 are now required to assess their students' progress against National Standards and this is new for us as a nation. Our primary and intermediate teachers will need to take part in ongoing moderation conversations as part of this requirement. The reason given for introducing the standards was to raise the achievement of all students, but most specifically those who might otherwise be at risk of falling behind their peers. Therefore, it is imperative for us to fill in the "black box" (Black & Wiliam, 1998) between the standards, which are after all simply a means to making a judgement about student work, and making substantive lifts in actual student achievement. Consequently, we undertook this review with the aim of finding out more about how teachers' improved knowledge and confidence about making assessment judgements might subsequently lead to richer learning opportunities, and hence achievement gains for students.

As we had anticipated, there are tantalising indications of the potential for powerful professional learning to emerge during moderation conversations. However, it will be evident that we found very little work that directly builds a clear chain of events and consequences. Consequently, we have begun to build the case piecemeal, with evidence drawn from here and there. This suggests to us that the area of moderation *as* professional learning is ripe for further research and development. Given the time our primary teachers are now destined to spend in moderation conversations, finding ways to make these rich professional learning conversations would appear to be no less than an urgent moral imperative (Fullan, 2010).

References

- Absolum, M., Flockton, L., Hattie, J., Hipkins, R., & Reid, I. (2009). *Directions for Assessment in New Zealand (DANZ)*. Wellington: Ministry of Education.
- Adie, L. E. (2008, November). *Changing assessment practices: The case for online moderation*. Paper presented at the Australian Association for Research in Education 2008 International Education Research conference, Brisbane. Retrieved from <http://www.aare.edu.au/08pap/adi08173.pdf>
http://ocs.sfu.ca/aare/index.php/AARE_2008/AARE/paper/viewFile/173/89
- Barrett, J. (2008). *The impact of the NCEA on teacher collegiality*. Unpublished PhD thesis, University of Waikato, Hamilton.
- Black, P., & Wiliam, D. (1998). *Inside the black box. Raising standards through classroom assessment*. London: nferNelson.
- Coburn, C. E., Toure, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record*, 111(4), 1115–1161.
- Duncan, C. R., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research*, 53(1), 1–21.
- Fazio, X., & Gallagher, T. L. (2009). Supporting learning: An examination of two teacher development collectives. [Article]. *Complicity: An International Journal of Complexity & Education*, 6(2), 1–19.
- Fullan, M. (2010). *All systems go: The change imperative for whole system reform*. Thousand Oaks, London, New Delhi, Singapore: Corwin/Sage.
- Gill, M. G., & Hoffman, B. (2009). Shared planning time: A novel context for studying teachers' discourse and beliefs about learning and instruction. *Teachers College Record*, 111(5), 1242–1273.
- Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. London: Falmer Press.
- Healy, H., & Bush, H. (2010). Moderation: Making learning a priority in primary religious education. *Journal of Religious Education*, 58(1), 29–37.
- Hipkins, R. (2010a). *The evolving NCEA*. Wellington: New Zealand Council for Educational Research.
- Hipkins, R. (2010b). Learning through moderation: Minding our language. *set: Research Information for Teachers*, 1, 18–19.
- Hipkins, R. (2010c). Reflections on being “labelled” by National Standards. *set: Research Information for Teachers*, 3, 27–30.
- Hipkins, R. (2011, May). *Learning to reinvent the school curriculum*. Paper presented at the Connected and Contagious: Exploring Learning Organisation Ideas in Schools, Wellington. conference. Retrieved from <http://learningemergence.net/wp-content/uploads/2011/05/Hipkins-2011-Learning-to-reinvent-the-school-curriculum-Final-May-24.pdf>
- Hunter, K., & Docherty, P. (2011). Reducing variation in the assessment of student writing. *Assessment and Evaluation in Higher Education*, 36(1), 109–124.
- Klenowski, V., & Adie, L. (2009). Moderation as judgement practice: Reconciling system level accountability and local level practice. *Curriculum Perspectives*, 29(1), 10–28.
- Klenowski, V., & Wyatt-Smith, C. (2010). Standards-driven reform Years 1–10: Moderation an optional extra? *Australian Educational Researcher*, 37(2), 21–39.

- Kuzich, S., Groves, R., O'Hare, S., & Pelliccione, L. (2010, November). *Building team capacity: Sustaining quality in assessment and moderation practices in a fully online unit*. Paper presented at the Australian Technology Network Assessment conference, Sydney.
- Limbrick, L., & Knight, N. (2005). Close reading of students' writing: What teachers learn about writing. *English Teaching: Practice and Critique*, 4(2), 5–22.
- Maxwell, G. S. (2002). *Moderation of teacher judgments in student assessment: Discussion paper*. Brisbane: Queensland School Curriculum Council.
- Maxwell, G. S. (2008, September). *Setting standards: Fitting form to function*. Paper presented at the International Association for Educational Assessment annual conference, Cambridge.
- McMillan, J. H., & Nash, S. (2000, April). *Teacher classroom assessment and grading practices decision making*. Paper presented at the National Council on Measurement in Education annual meeting, New Orleans.
- Millwood, K. (2007, December). *What impact does cross-curricular, collaborative moderation of asTTle writing tests have on student achievement in science?* Paper presented at the New Zealand Association for Research in Education conference, Christchurch.
- Ministry of Education. (2009a). *Mathematics standards for years 1–8*. Wellington: Learning Media.
- Ministry of Education. (2009b). *Reading and writing standards for years 1–8*. Wellington: Learning Media.
- Ministry of Education. (2010). *National standards factsheet*. Wellington: Author.
- Murphy, P., & Hall, K. (2008). *Learning and practice: Agency and identities*. London: Sage.
- Nixon, R., & McClay, J. K. (2007). Collaborative writing assessment: Sowing seeds for transformational adult learning. *Assessing Writing*, 12(2), 149–166.
- Parr, J. M., & Timperley, H. S. (2009). Teachers, schools and using evidence: Considerations of preparedness. *Assessment in Education: Principles, Policy and Practice*, 15(1), 57–71.
- Reid, L. (2007). Teachers talking about writing assessment: Valuable professional learning? *Improving Schools*, 10(2), 132–149.
- Scarino, A. (2005). Going beyond criteria for judging performance. *Babel*, 39(3), 8–38.
- Strachan, J. (1996, December). *Moderation: Swan or ugly duckling?* Paper presented at the New Zealand Association for Research in Education conference, Nelson.
- Timperley, H. S. (2008). *Teacher professional learning and development*. Belley, France: International Academy of Education.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy and Practice*, 14(3), 281–294.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy and Practice*, 17(1), 59–75.
- Wylie, C., & Hodgen, E. (2010). *NZCER 2010 primary and intermediate schools national survey: A snapshot of overall patterns and findings related to national standards*. Wellington: New Zealand Council for Educational Research.